

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 09:25:40 ; Search time 69 Seconds
(without alignments)
411.339 Million cell updates/sec

Title: US-09-896-580a-1

Perfect score: 1112
Sequence: 1 MTEYLNLFVTVQLKQIKIR.....KDHPLQHPDTAVEVQHRRH 213

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A: Geneseq 101002.*
1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID32/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1112	100.0	213	AA48343	Staphylococcus aur
2	938	84.4	183	AA483186	Staphylococcus aur
3	817	73.5	210	ABP38601	Staphylococcus epi
4	795	71.5	183	AA381798	S. epidermidis ope
5	754	67.8	150	AA483187	Staphylococcus aur
6	545	49.0	146	AA382977	S. epidermidis ope
7	533.5	48.0	183	AB47633	Listeria monocytog
8	477	42.9	204	ABP26597	Streptococcus poly
9	464	41.7	204	ABP26596	Streptococcus poly
10	457.5	41.1	203	AA477215	Streptococcus pneu

11	436	39.2	211	23	AB53869
12	386.5	34.8	155	19	AA486043
13	386.5	34.8	155	19	AA486043
14	230	20.7	216	23	AA48347
15	211.5	19.0	169	23	AA48345
16	206	18.5	273	21	AA48345
17	198	17.8	181	20	AA48345
18	197	17.7	158	21	AA48345
19	197	17.7	158	21	AA48345
20	186.5	16.8	160	20	AA48345
21	184.5	16.6	160	20	AA48345
22	184	16.5	160	20	AA48345
23	184	16.5	160	20	AA48345
24	184	16.5	160	20	AA48345
25	184	16.5	160	20	AA48345
26	180.5	16.2	168	23	AA48344
27	180.5	16.2	169	19	AA48344
28	179	16.1	170	23	AA48344
29	178	16.0	204	20	AA48344
30	176	15.8	166	20	AA48344
31	164	14.7	403	22	AB62406
32	160.5	14.4	193	22	AA48344
33	160.5	14.4	193	23	AA48344
34	158	14.2	155	21	AA48344
35	158	14.2	155	21	AA48344
36	152	13.7	136	23	ABP28421
37	151.5	13.6	174	19	AA48344
38	147	13.2	163	23	ABP39983
39	143	12.9	169	22	AA48344
40	141.5	12.7	214	22	AA48344
41	141	12.7	131	19	AA48344
42	141	12.7	136	19	AA48344
43	140	12.6	136	23	ABP28422
44	140	12.6	243	22	AA48344
45	140	12.6	243	23	AA48344

ALIGNMENTS

RESULT 1

AA48343
ID AA48343 standard; protein; 213 AA.

AC AA48343;

DT 23-APR-2002 (first entry)

XX Staphylococcus aureus peptide deformylase with C-terminal 6xHis tag.

DE Staphylococcus aureus peptide deformylase with C-terminal 6xHis tag.

XX Peptide deformylase; enzyme; protein coordinate data; antibacterial; pdf.

OS Staphylococcus aureus.

FN W0200202758-A2.

PD 10-JAN-2002.

XX 29-JUN-2001; 2001WO-US20777.

PR 30-JUN-2000; 2000US-215550P.

XX (PAA) PHARMACIA & UPJOHN CO.

XX Baldwin ET, Harris MS;

XX WPI; 2002-148012/19.

PT Crystalline Staphylococcus aureus peptide deformylase useful for solving structures of other molecules or molecular complexes, and designing modifiers of peptide deformylase activity

PS Claim 67; Fig 3; 149pp; English.

Lactococcus lactis
S. pneumoniae deri
Streptococcus pneu
Mycoplasma pneumon
Haemophilus influe
Arabidopsis thalia
Chlamydia trachoma
Arabidopsis thalia
Arabidopsis thalia
A formate transpor
Bacillus subtilis
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Escherichia coli p
E.coli peptide def
Staphylococcus aur
C. pneumoniae prot
Protein involved i
Drosophila melanog
C glutamicum prote
Corynebacterium g1
Arabidopsis thalia
Arabidopsis thalia
Streptococcus poly
H. pylori GHPO 266
C glutamicum prote
Propionibacterium
Streptococcus pneu
Streptococcus pneu
Streptococcus poly
Human mitochondria
Human peptide defo

XX The present sequence is Staphylococcus aureus peptide deformylase (pdf).
 CC The present invention relates to the crystal structure for pdf, which can
 CC be used in combination with a computer-assisted method for identifying,
 CC designing and making a potential modifier of S.aureus pdf activity.
 CC Modifiers of pdf are useful for blocking bacterial growth.
 XX
 SQ Sequence 213 AA;
 Query Match 100.0%; Score 1112; DB 23; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2.2e-113;
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MYEYLNLFVTIQLKQIRKQVQMLTMDIIRDGHPTLRQKAAELPLTKKEKETLIA 60
 Db 1 MYEYLNLFVTIQLKQIRKQVQMLTMDIIRDGHPTLRQKAAELPLTKKEKETLIA 60
 QY 61 MRFLVNSODEETAKYGLRSGVGLAARPOINISKRMIIVLPDGGSGKSYDMLNPKIV 120
 Db 61 MRFLVNSODEETAKYGLRSGVGLAARPOINISKRMIIVLPDGGSGKSYDMLNPKIV 120
 QY 121 SHSVQAYLPTGEGCLSDVDNVAAGLVHRHNKTIKAKOIEGNDIQLRLKGYPAIVFOHEI 180
 Db 121 SHSVQAYLPTGEGCLSDVDNVAAGLVHRHNKTIKAKOIEGNDIQLRLKGYPAIVFOHEI 180
 QY 181 DHLNGVMFYDHDKOHPLOPHTDAVEVHQHHH 213
 Db 181 DHLNGVMFYDHDKOHPLOPHTDAVEVHQHHH 213
 RESULT 2
 AA083186
 ID AA083186 standard; Protein; 183 AA.
 XX
 AC AA083186;
 XX
 DT 11-FEB-1999 (first entry)
 XX
 DE Staphylococcus aureus deformylase 1.
 XX
 DE Staphylococcus aureus; Defl; deformylase 1; bacterial infections;
 KW antimicrobial screening; diagnosis; antibacterial.
 XX
 OS Staphylococcus aureus.
 XX
 XX
 XX
 PD 25-NOV-1998.
 XX
 PF 21-MAY-1998; 98EP-0304076.
 XX
 PR 15-AUG-1997; 97US-0911844.
 PR 21-MAY-1997; 97US-0048706.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 XX Lonetto MA, Sylvestre DR, Warren RL;
 XX
 DR WPI; 1998-596871/51.
 DR N-PSDB; AAV70274.
 XX
 XX New nucleic acid encoding polypeptide deformylase of Staphylococcus
 PT aureus - useful for treatment prevention and diagnosis of bacterial
 PT infections, and for antimicrobial screening
 XX
 PS Claim 12; Page 19-20; 23pp; English.
 XX
 CC The present sequence represents deformylase 1 (Defl) isolated from
 CC Staphylococcus aureus WCH 29 (NCIMB 40771). Host cells containing
 CC vectors comprising nucleic acid molecules encoding Defl are used to
 CC produce recombinant Defl which have polypeptide deformylase (Defl)
 CC activity. Defl proteins are used to treat conditions requiring Defl
 CC activity, while Defl antagonists, e.g. Ab, are used to treat conditions

CC requiring reduction in Defl activity, especially they are antibacterials
 CC for treating a wide range of infections caused by Staphylococcus, also
 CC by some other bacteria, e.g. Helicobacter pylori. Defl proteins are also
 CC used to screen for compounds that interact specifically with it (i.e.
 CC potential antibacterials); to study the role of Defl in disease; for
 CC generating Ab and in protective vaccines (to generate an antibody and/or
 CC T-cell response). Vaccination may also be with a vector that contains a
 CC nucleic acid molecule encoding Defl. Ab are used therapeutically and for
 CC diagnosing in standard immunoassays. Antagonists may inhibit binding of
 CC bacteria to extracellular matrix proteins and to in-dwelling devices, or
 CC they inhibit normal progression of infection. Fragments of nucleic acid
 CC molecules encoding Defl are useful as hybridisation probes and/or
 CC amplification primers for isolation of full-length Defl-encoding
 CC sequences or related genes; to detect expression in infected tissue,
 CC e.g. for diagnosis and staging, and to identify mutations (e.g. for
 CC serotyping).
 XX
 SQ Sequence 183 AA;
 Query Match 84.4%; Score 938; DB 19; Length 183;
 Best Local Similarity 98.9%; Pred. No. 2e-94;
 Matches 181; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 25 MLTMDIIRDGHPTLRQKAAELPLTKKEKETLIAAREFLVNSODEETAKYGLRSGV 84
 Db 1 MLTMDIIRDGHPTLRQKAAELPLTKKEKETLIAAREFLVNSODEETAKYGLRSGV 60
 QY 85 LAAPQINISKRMIIVLPDGGSGKSYDMLNPKIVSHSVQAYLPTGEGCLSDVDNVA 144
 Db 61 LAAPQINISKRMIIVLPDGGSGKSYDMLNPKIVSHSVQAYLPTGEGCLSDVDNVA 120
 QY 145 LVHRHNKTIKAKOIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHDKOHPLOHTDA 204
 Db 121 LVHRHNKTIKAKOIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFYDHDKOHPLOHTDA 180
 QY 205 VEV 207
 Db 181 VEV 183
 RESULT 3
 ABP38601
 ID ABP38601 standard; Protein; 210 AA.
 XX
 AC ABP38601;
 XX
 DT 24-JUL-2002 (first entry)
 XX
 DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3446.
 XX
 KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.
 XX
 OS Staphylococcus epidermidis.
 XX
 XX US6380370-B1.
 XX
 PD 30-APR-2002.
 XX
 PF 13-AUG-1998; 98US-0134001.
 XX
 PR 14-AUG-1997; 97US-055779P.
 PR 08-NOV-1997; 97US-064964P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA
 XX Doucette-Stamm LA, Bush D;
 PI
 XX WPI; 2002-381255/41.
 DR N-PSDB; ABN91146.
 XX
 XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
 PT polypeptide, useful for diagnosing and treating bacterial infections -

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AAH81454 to AAH93120, from *Staphylococcus epidermidis*. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the *S. epidermidis* polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to AAH55090 represent specifically claimed *S. epidermidis* genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention.

N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO.4454 so even though sequences are given in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 183 AA;
XX

CC	sequences given in the sequence listing of the present specification,
CC	however the sequence listing only goes up to SEQ ID NO:4434 so even
CC	though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC	no sequences are present for SEQ ID NO:4455 to 4464.
XX	
SEQ	Sequence 183 AA;
	Query Match 71.5%; Score 795; DB 22; Length 183;
	Best local Similarity 79.8%; Pred. No. 9.6e-79;
	Matches 146; Conservative 23; Mismatches 14; Indels 0; Gaps 0;
QY	25 MLTWKDIIRDGHTPIROKAAEELSLPTKEEKETLIAMREFIVNSODEETAKYGIILRSVG 84
DB	1 MITWKKIIRDGHTPIREKAKELSLFPLSNNDKTFAMREFLINSODEETAKYGIILRSVG 60
QY	85 LAAPQINISKRMIAVLIPDDGSKSYDMLNPKIVSHSVQEAFLPTGGCLSLVDNVA 144
DB	61 LAAPQINEKRMIAVYLPDDGSKSYDMLNPKIMSYSVQEAFLPTGGCLSVDEITPG 120
QY	145 LVRRHNKHTIRAKOIEGNDIQLRLKGYPAIFQHEIDHLANGVMFYDHIKDHPHQFTDA 204
DB	121 LVRRHVRVTKAQDIDGNDVKRLKGYPAIFQHEIDHLANGIMFYDIDANEPFLKPHEEA 180
QY	205 VEV 207
DB	181 VEV 183

QY	205 VEV 207	
Db	181 VEV 183	
RESULT 5		
	AAW83187	
ID	AAW83187 standard; Protein; 150 AA.	
XX	AAW83187;	
XX	XX	
DT	11-FEB-1999 (first entry)	
XX	XX	
DE	Staphylococcus aureus deformylase 1 ORF protein sequence.	
XX	XX	
KN	Staphylococcus aureus; Defi; deformylase 1; bacterial infections;	
KW	antimicrobial screening; diagnosis; antibacterial.	
XX	XX	
OS	Staphylococcus aureus.	
XX	XX	
PN	EP879879-A2.	
XX	XX	
PD	25-NOV-1998.	
XX	XX	
PF	21-MAY-1998; 98EP-0304076.	
XX	XX	
PR	15-AUG-1997; 97US-0911844.	
PR	21-MAY-1997; 97US-0048706.	
XX	XX	
PA	(SMIK) SMITHKLINE BEECHAM CORP.	
XX	XX	
PI	Lonetto WA, Sylvester DR, Warren RL;	
XX	XX	
DR	WPI; 1998-596871/51.	

DR N-PSDB; AAV70275.
 XX
 PT New nucleic acid encoding polypeptide deformylase of *Staphylococcus aureus* - useful for treatment, prevention and diagnosis of bacterial infections, and for antimicrobial screening
 PT
 XX
 PS Disclosure; Page 21; 23pp; English.
 XX
 XX The present sequence represents deformylase 1 (Def1) isolated from *Staphylococcus aureus* WCUH 29 (NCIMB 40771). Host cells containing vectors comprising nucleic acid molecules encoding Def1 are used to produce recombinant Def1 which have polypeptide deformylase (Def) activity. Def1 proteins are used to treat conditions requiring Def activity, while Def1 antagonists, e.g. Ab, are used to treat conditions requiring reduction in Def activity, especially they are antibacterials for treating a wide range of infections caused by *Staphylococcus*, also by some other bacteria, e.g. *Helicobacter pylori*. Def1 proteins are also used to screen for compounds that interact specifically with it (i.e. potential antibacterials); to study the role of Def in disease; for generating Ab and in protective vaccines (to generate an antibody and/or T-cell response). Vaccination may also be with a vector that contains a nucleic acid molecule encoding Def1. Ab are used therapeutically and for diagnosing in standard immunoassays. Antagonists may inhibit binding of bacteria to extracellular matrix proteins and to in-dwelling devices, or they inhibit normal progression of infection. Fragments of nucleic acid molecules encoding Def1 are useful as hybridisation probes and/or amplification primers for isolation of full-length Def-encoding sequences or related genes; to detect expression in infected tissue, e.g. for diagnosis and staging, and to identify mutations (e.g. for serotyping).
 XX
 SQ Sequence 150 AA;

Query Match 67.8%; Score 754; DB 19; Length 150;
 Best Local Similarity 98.7%; Pred. No. 2.2e-74;
 Matches 148; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 25 MLTKDIIIRGHTLRQAAELPLTKKEKTLIAMREFLVNSQDEEIAKRYGLRSGVG 84
 Db 1 MLTKDIIIRGHTLRQAAELPLTKKEKTLIAMREFLVNSQDEEIAKRYGLRSGVG 60
 QY 85 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPKIVSHSVQEAFLPTGEGCLSVDDNVAG 144
 Db 61 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPKIVSHSVQEAFLPTGEGCLSVDDNVAG 120
 QY 145 LVHRNKITKADIEGNDIOIRLKGYPAI 174
 Db 121 LVHRNRITKADIEGNDIOIRLKGYPAM 150

RESULT 6
 AAG82977
 ID AAG82977 standard; Protein; 146 AA.

XX
 AC AAG82977;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:3048.
 XX
 KW *Staphylococcus epidermidis* SRI strain; infection; diagnosis;
 KW vaccination; endocarditis.
 XX
 OS *Staphylococcus epidermidis*.
 XX
 PN WO200134809-A2.
 XX
 XX 17-MAY-2001.
 XX
 XX 09-NOV-2000; 2000WO-US30782.
 XX
 XX 09-NOV-1999; 99US-0164258.
 XX

(GLAX) GLAXO GROUP LTD.
 XX
 PI Kimmerly WU;
 XX
 DR WPI; 2001-316495/33.
 DR N-PSDB; AAH53827.
 XX
 XX Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*, useful for vaccinating against infections, e.g. endocarditis -
 PT
 PT Claim 18; Page 802; 2188pp; English.
 XX
 PS
 XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*.
 CC (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the S. epidermidis polypeptides (II) via the production of vectors containing them which are used to produce host cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to AAH55090 represent specifically claimed S. epidermidis genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though the sequences are given in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 SQ Sequence 146 AA;

Query Match 49.0%; Score 545; DB 22; Length 146;
 Best Local Similarity 81.7%; Pred. No. 1.7e-51;
 Matches 103; Conservative 12; Mismatches 11; Indels 0; Gaps 0;
 QY 25 MLTKDIIIRGHTLRQAAELPLTKKEKTLIAMREFLVNSQDEEIAKRYGLRSGVG 84
 Db 1 MLTKDIIIRGHTLRQAAELPLTKKEKTLIAMREFLVNSQDEEIAKRYGLRSGVG 60
 QY 85 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPKIVSHSVQEAFLPTGEGCLSVDDNVAG 144
 Db 61 LAAPQINISKRMIAVLIPDDGSGKSYDMLVNPKIVSHSVQEAFLPTGEGCLSVDDNVAG 120
 QY 145 LVHRNH 150
 Db 121 LVHRHH 126

RESULT 7
 ABB47633
 ID ABB47633 standard; Protein; 183 AA.

XX
 AC ABB47633;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE *Listeria monocytogenes* protein #337.
 XX
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 XX
 OS *Listeria monocytogenes*.
 XX
 XX WO200177335-A2.
 XX
 XX 18-OCT-2001.
 XX
 XX 11-APR-2001; 2001WO-FR01118.
 XX

```
PR 11-APR-2000; 2000FR-0004629.
XX (INSP ) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Medjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Hierrez-Martinez A, Amend A;
PI Chakraborty T, Domann B, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides -
XX
XX Claim 6; SEQ ID NO 338; 192pp; French.
XX
XX The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 183 AA;
Query Match 48.0%; Score 533.5; DB 23; Length 183;
Best Local Similarity 57.5%; Pred. No. 4.4e-50;
Matches 104; Conservative 30; Mismatches 44; Indels 3; Gaps 2;
QY 25 MLTWKDIIRDCGHTPLROKAALELPYKEKETLIAMREFLVNSQDEEIAKRYGIRSGVG 84
Db 1 MLTMDIIVREGHPALREAVATEVTPLSDEEKKLGDRMDLEFLINSQDEDLAEKRYGLRGVG 60
QY 85 LAAPQINISKRMIAVLIIPDGGSGKSYDMLVNPKVSHSVQVQAYLPTGEGCLSVDDNVAG 144
Db 61 IAAPQLAVTKRFLAIHVHDE-KDLNLSYVLNPKIRSHSVQACLSGEGCLSVDRVEFG 119
QY 145 LVVRHNKTIKAKDIEGNDIOLRLKGYPAIVFOHEIDHNGVMFYDHIKQHP-LQPHF 202
Db 120 YVRSERVITADFENGTPKLRFKGYPAIVIOHEIDHNGIMFYDHNKENSYPYDPV 179
QY 203 D 203
Db 180 D 180
RESULT 8
ABP26597
ID ABP26597 standard; Protein; 204 AA.
XX
XX AC ABP26597;
XX
XX DT 02-JUL-2002 (first entry)
XX
XX DE Streptococcus polypeptide SEQ ID NO 2370.
XX
```

```
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
XX W0200234771-A2.
XX
XX 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB04789.
XX
XX 27-OCT-2000; 2000GB-0026333.
XX
XX 24-NOV-2000; 2000GB-0028727.
XX
XX 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
XX
XX (GENO-) INST GENOMIC RES.
XX
XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
XX Tettelin H;
XX
XX WPI; 2002-352536/38.
XX
XX N-PSDE; ABN67228.
XX
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
XX Claim 1; Page 3389; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and anti-inflammatory
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
XX SQ Sequence 204 AA;
Query Match 42.9%; Score 477; DB 23; Length 204;
Best Local Similarity 52.6%; Pred. No. 8.1e-44;
Matches 101; Conservative 33; Mismatches 50; Indels 8; Gaps 2;
QY 16 QTKIRKQVMTKMDIIRDCGHTPLROKAALELPYKEKETLIAMREFLVNSQDEEIAK 75
Db 4 QDKLIKPSHLITMDIIRDCGHTPLROKAALELPYKEKETLIAMREFLVNSQDEEIAK 63
QY 76 RYGLSGVGLAAPQINISKRMIAVLIIPD--DGSCK-----SYDVMVNPKVSHSVQEA 127
Db 64 KLGLRAGVGLAAPQINISKRMIAVLIIPDKNPPKPKYKAYSWQEVLPNPKVSHSVQDA 123
QY 128 YLPTGEGCLSVDDNVAGLVVRHNKTIKAKDIEGNDIOLRLKGYPAIVFOHEIDHNGVM 187
Db 124 ALSDGECLSVDRVREGVVRHARVTVYDKSQOQRHKLKGNALVIVQHEIDHNGVL 183
QY 188 FYDHIDKHPLQ 199
Db 184 FYDRINAKNPF 195
RESULT 9
```

ABP26596

ID ABP26596 standard; Protein; 204 AA.

XX AC ABP26596;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 2368.

XX KW Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
 XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 XX KW antinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus agalactiae.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB04789.

XX PR 27-OCT-2000; 2000GB-0026333.

XX PR 24-NOV-2000; 2000GB-0028727.

XX PR 07-MAR-2001; 2001GB-0005640.

XX PA (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;

XX PI Tettelin H;

XX DR WPI; 2002-352536/38.

XX DR N-PSDB; ABN67227.

XX FS Claim 1; Page 3389; 4525pp; English.

XX CC The invention relates to a protein (ABP25413-ABP30895) from group B streptococcus (GBS (Streptococcus agalactiae) or group A streptococcus (GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and antiinflammatory activity. (I), nucleic acids encoding (I), ABN65044-ABN71526 and antibodies that bind (I) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. Nucleic acids encoding (I) are used to detect Streptococcus in a biological sample. (I) is used to determine whether a compound binds to (I). A composition comprising (I) or a nucleic acid encoding (I) may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (I) may be used to recombinantly produce (I) and may be used in gene therapy. Antibodies to (I) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins.

XX SQ Sequence 204 AA;

Query Match 41.7%; Score 464; DB 23; Length 204;

Best Local Similarity 50.5%; Pred. No. 2.2e-42;

Matches 95; Conservative 38; Mismatches 47; Indels 8; Gaps 2;

QY 18 KIRKVOYMLTKWDIRGHPTRQKAAELPLTKKEKETLIAMREFLYNSQDEIAKEY 77

DB 6 KLVKASHIDNDIIRGNPALTVAEVTPLPSQDEILIGERKMOFLKHSQDPVNAEKM 65

QY 78 GLRSGVGLAAPQINISKRIMAVLIP-----SYDMLVWPKIVSHSVQEAYL 129

DB 66 GLRSGVGLAAPQIDISKRIITAVLVNVEVDAGNPPKAYSLQEVWYNNPKVSHSVQDAAL 125

QY 130 PTGEGCLSVDDNVAGLVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFY 189

DB 126 SDGEGCLSVDEVPVGYVHRHATVIEYFDKGTGKXRLKLGKYNISIVQGEIDHIDGIMFY 185
 QY 190 DHIDKXDP 197
 DB 186 DRINEKNP 193

RESULT 10

AAW77215

ID AAW77215 standard; Protein; 203 AA.

XX AC AAW77215;

XX DT 16-NOV-1998 (first entry)

XX DE Streptococcus pneumoniae defl polypeptide sequence.

XX KW Streptococcus pneumoniae; otitis media; conjunctivitis; meningitis;
 XX KW pneumonia; bacteraemia; sinusitis; pleural empyaema; endocarditis; Defl.

XX OS Streptococcus pneumoniae.

XX PH Key Location/Qualifiers

XX FT Modified-site 26 /note= "Xaa is unspecified, encoded by WCT"

XX PN EP863152-A2.

XX PD 09-SEP-1998.

XX PF 05-FEB-1998; 98EP-0300873.

XX PR 15-AUG-1997; 97US-0911503.

XX PR 10-FEB-1997; 97US-0037536.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PA (SMIK) SMITHKLINE BEECHAM PLC.

XX PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;

XX PI Stodola RK;

XX DR WPI; 1998-458798/40.

XX DR N-PSDB; AAV48308.

XX CC New DNA encoding pneumococcal deformylase polypeptides - used to prevent, treat, diagnose and vaccinate against e.g. otitis media, conjunctivitis, pneumonia, bacteraemia, meningitis, sinusitis, pleural empyaema and endocarditis

XX PS Claim 14; Page 5; 25pp; English.

XX CC The Streptococcus pneumoniae Defl protein can be used in the treatment, prevention and diagnosis of and vaccination against otitis media, conjunctivitis, pneumonia, bacteraemia, meningitis, sinusitis, pleural empyaema and endocarditis.

XX SQ Sequence 203 AA;

Query Match 41.1%; Score 457.5; DB 19; Length 203;

Best Local Similarity 50.3%; Pred. No. 1.1e-41;

Matches 94; Conservative 37; Mismatches 49; Indels 7; Gaps 2;

QY 18 KIRKVOYMLTKWDIRGHPTRQKAAELPLTKKEKETLIAMREFLYNSQDEIAKEY 77

DB 6 RTKAAHLIDNDIIRGNPALTVAEVTPLPSQDEILIGERKMOFLKHSQDPVNAEKM 65

QY 78 GLRSGVGLAAPQINISKRIMAVLIP-----DGSGKSYDY--MLVNPKIVSHSVQEAYL 130

DB 66 GLRSGVGLAAPQIDISKRIITAVLVNVEVDAGNPPKAYSLQEVWYNNPKVSHSVQDAAL 125

QY 131 TGEGLCLSVDDNVAGLVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFOHEIDLHNGVMFY 190

Db 126 EGGGLSDVRNPGYVWRHARVTVDFDKDGEKRIKLGYSIWOHEIDHNGIMFYD 185
 QY 191 HIKDHP 197
 Db 186 RINEKDP 192

RESULT 11
 ABB53869
 ID ABB53869 standard; Protein; 211 AA.

XX AC ABB53869;
 XX 16-MAY-2002 (first entry)
 DT Lactococcus lactis protein def.
 DE
 XX
 XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX Lactococcus lactis IL1403.

XX FR2807446-A1.

XX 12-OCT-2001.

XX 11-APR-2000; 2000FR-0004630.

XX 11-APR-2000; 2000FR-0004630.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;

XX WPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification of Lactococcus
 PT lactis and related species -

XX Claim 6; SEQ ID No 571; 2504pp; French.

XX The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (AA90521) and related proteins (AB553300-AB553621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 211 AA;

Query Match 39.2%; Score 436; DB 23; Length 211;
 Best Local Similarity 47.9%; Pred. No. 2.7e-39;
 Matches 93; Conservative 38; Mismatches 51; Indels 12; Gaps 3;

QY 16 QIKIRVQVLMJWKDIIRDGHTLROKAELELPITKKEKTLTAMERIVNSQDEIAK 75

Db 7 QANLIZKASHWISNDLIREGFTLREVRANDVLPDSDIEDILGKMLQFLHNSQDPWMAE 66

QY 76 RYGLRSGVGLAAPQINISKRMIAVLIP-----DGKSYDY--MLVNPKIVSHS 123

Db 67 KMWGLRGVGLAANQLGLLKKVIAVLIPNEPEVDEDGNETPPKEAYKMEIMYNAKVSHS 126

QY 124 VQRAYLPTGEGCLSDVDNAGLVHRHNKTIITAKDIEGNDIQLRKGYPAPVQHEIDL 183

Db 127 VQDAVVEGEGCLSDREVPGYVVRHARVTVETYNKEGKKIRLKDPFPAICVQHEIDT 186

QY 184 NGVMFYDHDIDKDP 197

Db 187 NGVMFYDHIWMDP 200

RESULT 12

AA86043

ID AAY86043 standard; Protein; 155 AA.

XX AC AAY86043;

DT 10-APR-2000 (first entry)

XX S. pneumoniae derived protein #252.

XX Treatment; prevention; disease; diagnosis; gene therapy; screening;
 KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.

XX Streptococcus pneumoniae.

XX WO9806734-A1.

XX 19-FEB-1998.

XX 15-AUG-1997; 97WO-US14436.

XX 16-AUG-1996; 96US-0024022.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
 PI Stodola RK;

XX WPI; 1998-159452/14.

XX N-PSDB; AA296374.

XX Streptococcus pneumoniae proteins and related DNA - useful for
 PT screening compounds for antibacterial activity

XX Claim 5; Page 528; 640pp; English.

XX This invention describes novel isolated Streptococcus pneumoniae
 CC polynucleotides (see AA296173-296494) and their encoded proteins (see
 CC AA295793-296182). The DNA, vectors and host cells described in the
 CC method of the invention are useful for the recombinant expression of the
 CC polypeptides. The polypeptides are useful for treatment or prevention of
 CC disease, or diagnosis of disease related to expression or activity of
 CC such a polypeptide. They can also be used to screen for compounds which
 CC interact with and inhibit or activate such a polypeptide. The
 CC polypeptides (or DNA encoding them, via gene therapy) are also useful
 CC for inducing an immunological response in a mammal. The antagonists are
 CC useful to inhibit such bacterial polypeptides. The polypeptides are
 CC particularly useful to identify antimicrobial compounds and antibiotics.
 CC They are also useful to determine their role in pathogenesis of
 CC infection, dysfunction and disease.

XX Sequence 155 AA;

Query Match 34.8%; Score 386.5; DB 19; Length 155;
 Best Local Similarity 54.2%; Pred. No. 4.4e-34;
 Matches 78; Conservative 28; Mismatches 31; Indels 7; Gaps 2;

QY 61 MREFLVNSQDEIAKRYGLRSGVGLAAPQINISKRMIAVLIPD-----DGKSKSYDY--M 113

Db 1 MMQFLKHSQDPWMAEKRLGGLGGLAAPQLDISKRIITAVLPVNIIVEGETPQEAYLEAI 60

QY 114 LVNPKIVSHSVQRAYLPTGEGCLSDVDNAGLVHRHNKTIITAKDIEGNDIQLRKGYP 173

Db 61 MYNPKIVSHSVQDAALGEGEGCLSDVDNPGVVRHARVTVDFDKDGEKRIKLGINS 120

QY 174 IVFQHEIDLHNGVMFYDHDIDKDP 197

Db 121 IVVQHEIDHNGIMFYDRIKNDP 144

RESULT 13

AAW77216
 ID AAW7216 standard; Protein; 155 AA.
 AC
 AC AAW7216;
 DT
 DT 16-NOV-1998 (first entry)
 XX
 DE Streptococcus pneumoniae defl open reading frame peotide sequence.
 XX
 KW Streptococcus pneumoniae; otitis media; conjunctivitis; meningitis;
 KW pneumonia; bacteraemia; sinusitis; pleural empyaema; endocarditis; Defl.
 XX
 OS Streptococcus pneumoniae.
 XX
 FN EP863152-A2.
 XX
 PD 09-SEP-1998.
 XX
 PF 05-FEB-1998; 98EP-0300873.
 XX
 PR 15-AUG-1997; 97US-0911503.
 PR 10-FEB-1997; 97US-0037536.
 XX
 PA (SMIX) SMITHKLINE BEECHAM CORP.
 PA (SMIX) SMITHKLINE BEECHAM PLC.
 PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
 PI Stodola RK;
 XX
 DR WPI; 1998-458798/40.
 DR N-PSDB; AAW48309.
 XX
 XX New DNA encoding pneumococcal deformylase polypeptides - used to
 PT prevent, treat, diagnose and vaccinate against e.g. otitis media,
 PT conjunctivitis, pneumonia, bacteraemia, meningitis, sinusitis,
 PT pleural empyaema and endocarditis
 XX
 PS Claim 23; Page 5; 25pp; English.
 XX
 CC The Streptococcus pneumoniae Defl protein can be used in the treatment,
 CC prevention and diagnosis of and vaccination against otitis media,
 CC conjunctivitis, pneumonia, bacteraemia, meningitis, sinusitis, pleural
 CC empyaema and endocarditis.
 XX
 SQ Sequence 155 AA;
 Query Match 34.8%; Score 386.5; DB 19; Length 155;
 Best Local Similarity 54.2%; Pred. No. 4.4e-34;
 Matches 78; Conservative 28; Mismatches 31; Indels 7; Gaps 2;
 QY 61 MREFLVNSODEETAKRYGLRSVGLAAPOLNISKRMIAVLIPD-----D3SGKSYDY-M 113
 DB 1 MMEFLKHSQDPVMAEKXGLRGVGLAAPQLDISKR-IIVLVNIVVEGETPQRAYDLAI 60
 QY 114 LVNPKIVSHSVQAYLPTGEGCLSDVDNAGLVHRHNKTIKAKDIEGNDIQLRLKGYPA 173
 DB 61 MYNPKIVSHSVQDAALGEGCLSDVRNVPVYVHRVTVDFDKDGEKRIKLGYS 120
 QY 174 IVFOHEIDHLNGVMFYDHDKDHP 197
 DB 121 IVVQHEIDHLNGIMFYDRINEKOP 144
 RESULT 14
 AAW48347
 ID AAW48347 standard; protein; 216 AA.
 AC
 AC AAW48347;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Mycoplasma pneumoniae peptide deformylase.

KW Peptide deformylase; enzyme; protein coordinate data; antibacterial; pdf.
 XX Mycoplasma pneumoniae.
 OS
 PN WO200202758-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 29-JUN-2001; 2001WO-US20777.
 XX
 PR 30-JUN-2000; 2000US-215550P.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Baldwin ET, Harris MS;
 XX
 DR WPI; 2002-148012/19.
 XX
 PT Crystalline Staphylococcus aureus peptide deformylase useful for
 PT solving structures of other molecules or molecular complexes, and
 PT designing modifiers of peptide deformylase activity -
 XX
 PS Example 1; Fig 3; 149pp; English.
 XX
 CC The present invention relates to the crystal structure for Staphylococcus
 CC aureus peptide deformylase (pdf; see AAW48343 for pdf protein sequence),
 CC which can be used in combination with a computer-assisted method for
 CC identifying, designing and making a potential modifier of S.aureus pdf
 CC activity. Modifiers of pdf are useful for blocking bacterial growth. The
 CC present sequence is Mycoplasma pneumoniae peptide deformylase (pdf),
 CC which was used in a sequence alignment with the S. aureus pdf sequence.
 XX
 SQ Sequence 216 AA;
 Query Match 20.7%; Score 230; DB 23; Length 216;
 Best Local Similarity 30.4%; Pred. No. 1e-16;
 Matches 59; Conservative 44; Mismatches 83; Indels 8; Gaps 5;
 QY 7 NLTVTIQ-LKQIKRKQVQMLTKDIIRGHPTLRQKAELELITKEEKETLIANREFL 65
 DB 11 SIFRIILLPQI---NMELLPTKAWLVLDVKEINPTKPVQFELDQASLDICAKMAYV 67
 QY 66 VNSODEETAKRYGLRSVGLAAPOLNISKRMIAVLIPDGSKSYDYMLVNPKIVSHSVQ 125
 DB 68 DASVNGD-AEKYIIPGIGIARAQIGVWQMFYIHLMD--CGVEHKLLINFLINLSAN 124
 QY 126 EAYLPTGEGCLSDVDNAGLVHRHNKTIKAKD-TEGNDIQLRLKGYPAIVFOHEIDHLN 184
 DB 125 KSPFKSGEGCLSVPRKHQGVYIRHEWITITGFDLQOKEITITATGLFGMCLQHEFDHLQ 184
 QY 185 GVMFYDHDKDHP 198
 DB 185 GRPYTHRIINPLNPL 198
 RESULT 15
 AAW48345
 ID AAW48345 standard; protein; 169 AA.
 XX
 AC AAW48345;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Haemophilus influenzae peptide deformylase.
 KW Peptide deformylase; enzyme; protein coordinate data; antibacterial; pdf.
 XX Haemophilus influenzae.
 OS
 PN WO200202758-A2.
 XX
 PD 10-JAN-2002.
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 09:34:30 ; Search time 26 Seconds
(without alignments)
241.042 Million cell updates/sec

Title: US-09-896-580A-1
Perfect score: 1112
Sequence: 1 MVEYLNLFVTLQKQIKR.....KDHPLQHTDAVEHQHHH 213

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	938	84.4	183	4	US-09-373-953-2
2	817	73.5	210	4	US-09-134-001C-3446
3	754	67.8	150	4	US-09-373-953-4
4	457.5	41.1	203	4	US-08-991-023-2
5	386.5	34.8	155	4	US-08-991-023-4
6	198	17.8	181	2	US-08-932-142-2
7	198	17.8	181	4	US-09-342-458-2
8	186.5	16.8	160	4	US-09-194-146-8
9	180.5	16.2	169	2	US-08-895-939-4
10	180.5	16.2	169	4	US-09-188-820-4
11	147	13.2	163	4	US-09-134-001C-4828
12	89	8.0	689	1	US-08-248-021A-2
13	84	7.6	2509	1	US-08-469-005A-10
14	84	7.6	2511	4	US-09-261-907-2
15	81	7.3	638	2	US-08-557-122A-38
16	81	7.3	638	4	US-09-262-666-38
17	81	7.3	1447	4	US-09-376-330-17
18	79.5	7.1	1375	4	US-09-722-139-2
19	79.5	7.1	1375	4	US-09-721-832-2
20	79.5	7.1	1375	4	US-09-721-689-2
21	78.5	7.1	600	3	US-09-212-971-12
22	78.5	7.1	600	4	US-08-800-929A-12
23	78.5	7.1	604	1	US-09-617-053A-12
24	78.5	7.1	604	4	US-08-487-753-2
25	78.5	7.1	604	2	US-08-480-065-2
26	78.5	7.1	604	3	US-08-487-744-2
27	78.5	7.1	604	5	PCT-US93-09167-2

28	76	6.8	625	1	US-08-242-689-2	Sequence 2, Appli
29	76	6.8	625	4	US-08-468-583-2	Sequence 2, Appli
30	75.5	6.8	605	3	US-08-693-214-8	Sequence 8, Appli
31	75	6.7	215	2	US-08-385-335A-12	Sequence 12, Appli
32	75	6.7	215	2	US-08-659-206A-5	Sequence 2, Appli
33	75	6.7	215	2	US-08-659-206A-5	Sequence 5, Appli
34	75	6.7	594	4	US-09-124-141-15	Sequence 15, Appli
35	75	6.7	594	4	US-09-124-141-23	Sequence 23, Appli
36	74.5	6.7	350	1	US-08-184-252A-2	Sequence 2, Appli
37	74.5	6.7	350	2	US-09-123-851-3	Sequence 3, Appli
38	74.5	6.7	350	2	US-08-728-520-3	Sequence 3, Appli
39	74.5	6.7	350	5	PCT-US95-00601-2	Sequence 2, Appli
40	74.5	6.7	1170	1	US-08-313-288B-20	Sequence 20, Appli
41	74	6.7	252	4	US-09-134-001C-5644	Sequence 5644, Ap
42	74	6.7	288	4	US-09-134-001C-3333	Sequence 3333, Ap
43	74	6.7	471	4	US-09-134-001C-3150	Sequence 3150, Ap
44	74	6.7	594	2	US-08-592-696-2	Sequence 2, Appli
45	74	6.7	594	2	US-09-027-536-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-373-953-2
; Sequence 2, Application US/09373953
; Patent No. 6410688
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A
; APPLICANT: Sylvester, Daniel
; APPLICANT: Warren, Richard
; TITLE OF INVENTION: No. 6410688el defL
; FILE REFERENCE: GM10001
; CURRENT APPLICATION NUMBER: US/09/373,953
; CURRENT FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: US 08/911,844
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: US 60/048,706
; PRIOR FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-373-953-2

Query Match	84.4%	Score	938	DB	4	Length	183
Best Local Similarity	98.9%	Pred. No.	1.2e-96				
Matches	181	Conservative	2	Mismatches	0	Indels	0
Gaps	0						
QY	25	MLTMKDIIIRDGHPTLRQKALELPITKKEKTLTAMFEFLVNSQDEETAKRYGLRSGVG	84				
Db	1	MLTMKDIIIRDGHPTLRQKALELPITKKEKTLTAMFEFLVNSQDEETAKRYGLRSGVG	60				
QY	85	LAAPQINISKRMIAVLIPDDGSKSYDYMVNPKIVSHSVQEAFLPTGEGCLSVDDNVAG	144				
Db	61	LAAPQINISKRMIAVLIPDDGSKSYDYMVNPKIVSHSVQEAFLPTGEGCLSVDDNVAG	120				
QY	145	LVRHNKTIITAKDIEGNDIQLKGYPAIVFQHIHDLNKGWMPFYDHDKDHPLOPHPTDA	204				
Db	121	LVRHNKTIITAKDIEGNDIQLKGYPAIVFQHIHDLNKGWMPFYDHDKDHPLOPHPTDA	180				
QY	205	VEV 207					
Db	181	VEV 183					

RESULT 2
US-09-134-001C-3446
; Sequence 3446, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:

Patent No. 6284878
GENERAL INFORMATION:
APPLICANT: Lonetto, Michael A.
TITLE OF INVENTION: def1
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,023
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM50010
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 155 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-991-023-4

Query Match 34.8%; Score 386.5; DB 4; Length 155;
Best Local Similarity 54.2%; Pred. No. 2.5e-35;
Matches 78; Conservative 28; Mismatches 31; Indels 7; Gaps 2;

QY 61 MRPELVNSODEETAKYGLRSGVGLAAPQINISKRMIAVLIPD-----DGSKSYDY--M 113
DB 1 MMQFLKHSQDPVNAERKMGRLGSGVGLAAPQLDISKRIIAVLVNPVVEGETPQRAYDLRAI 60
QY 114 LVNPKIVSHSVQAYLPTGEGCLSYDDNVAGLVHRHNKITIKAKDIEGNDIQLRLKGYP 173
DB 61 MYNPKIVSHSVQDAALGEGCLSYDRNVPGVYVVRHARVTVDFDKDGEKRIKLGKYN 120
QY 174 IVFQHEIDHLNGVMFYDHDKHP 197
DB 121 IVVQHEIDHNGIMFYDRINEKP 144

RESULT 6
US-08-932-142-2
Sequence 2, Application US/08932142
Patent No. 5962666
GENERAL INFORMATION:
APPLICANT: Lonetto, Michael A.
TITLE OF INVENTION: NOVEL def
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,142
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: GM10095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 181 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-932-142-2

Query Match 17.8%; Score 198; DB 2; Length 181;
Best Local Similarity 33.5%; Pred. No. 3.2e-14;
Matches 60; Conservative 31; Mismatches 62; Indels 26; Gaps 8;

QY 31 IIRD---GHPTLRKAAELELPLTKKEKETTIAARREFLVNSQDEEIAKRYGLRSGVGLA 86
DB 1 MIRDLEYDVSFILRKVAAPV-TEIDELRQLVLDMSFTM-----AFYKGVGLA 47
QY 87 APOINISKRMIAVLIP---DGSKSYDY--MLVNPKIVSHSVQAYLPTGEGCLSYDDN 141
DB 48 APQVGQSILFINGVERELEDGELVFCDFPRVFINPVITQKSEQLVY--GNEGCLSI-LR 104
QY 142 VAGLVHRHNKITIKAKDIEGNDIQLRLKGYPALVFPQHEIDHLNGVMFYDHI-DKDHPLQ 199
DB 105 IRGEVAPDKITESAKNLDQQLSLALEGFLARIWMHETDHLHGLVLYIDRMSDKDKTKQ 163

RESULT 7
US-09-342-458-2
Sequence 2, Application US/09342458
Patent No. 6307022
GENERAL INFORMATION:
APPLICANT: Lonetto, Michael A.
TITLE OF INVENTION: NOVEL def
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/342,458
FILING DATE: 29-Jun-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/932,142
FILING DATE: 16-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354

```
;
; REFERENCE/DOCKET NUMBER: GM10095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-342-458-2
;
;
; Query Match 17.8%; Score 198; DB 4; Length 181;
; Best Local Similarity 33.5%; Pred. No. 3.2e-14;
; Matches 60; Conservative 31; Mismatches 62; Indels 26; Gaps 8;
;
QY 31 IIRD-----GPTTQKAAELEPLTKKETLIAMREFLVNSQDEIAKRYGLSGVGLA 86
Db 1 MIRDLEYDSPILRKVAAPV-TEITDELRLQLVLDMSGT-----AFYKGVGLA 47
;
QY 87 APQINISKRMIAVLIP---DDSGSKSYD--MLVNPKIVSHSVQAYLPTGEGCLSVDDN 141
Db 48 APQVGQISLFIINGVERELEDGELVFCDFPRVFNPIVTKSEQLVY--GNEGCLSI-LR 104
;
QY 142 VAGLVHRNKITTKAKDIGNDIQRLKGYPAIVFOHEIDHLNGVMFYDHI-DKHPLQ 199
Db 105 IRGEVAPDPKITESAKNLDQQLSALGFLARVWHTDHLHGLVLYIDRMSDKTKQ 163
;
;
; RESULT 8
; US-09-194-146-8
; Sequence 8, Application US/09194146
; Patent No. 6458557
; GENERAL INFORMATION:
; APPLICANT: Miller, Brian
; APPLICANT: Diaz-Torres, Maria
; TITLE OF INVENTION: Gram-Positive Microorganism Formate Pathway
; FILE REFERENCE: GC395-US
; CURRENT APPLICATION NUMBER: US/09/194,146
; CURRENT FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: GB 9724627.6
; PRIOR FILING DATE: 1997-11-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Bacillus
; US-09-194-146-8
;
; Query Match 16.8%; Score 186.5; DB 4; Length 160;
; Best Local Similarity 35.7%; Pred. No. 5e-13;
; Matches 61; Conservative 26; Mismatches 59; Indels 25; Gaps 8;
;
QY 26 LTMKDLIDGHPIT-LRQKAAELEPLTKKETLIAMREFLVNSQDEIAKRYGLSGVG 84
Db 1 LAVKKVVT--HRAVLETFETPTVTVDKXKLLDDMDYDTMLE-----MDGVG 46
;
QY 85 LAAPQINISKRMIAVLIPDDSGSKSYDMLVNPKIVSHSVQAYLPTG-EGCLSVDDNVA 143
Db 47 LAAPQIGILKFAAVVEIGDD-RGR---IDLVPKLEKSGEQ-----TGEGCLSP-FNVY 97
;
QY 144 GLVHRNKITTKAKDIGNDIQRLKGYPAIVFOHEIDHLNGVMFYDHIK 194
Db 98 GDVTRADYVVKRAFNRQKPGFFILEARGFLAARVQHEMDHLDGVLFTSKISK 148
;
;
; RESULT 9
; US-08-895-939-4
; Sequence 4, Application US/08895939
; Patent No. 5834243
```

```
;
; GENERAL INFORMATION:
; APPLICANT: Bogosian, Gregg
; TITLE OF INVENTION: Deformylation of f-Met Peptides in
; Bacterial Expression Systems
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895,939
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,687
; FILING DATE: 17-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Patterson, Melinda L.
; REGISTRATION NUMBER: 33,062
; REFERENCE/DOCKET NUMBER: MOPV:006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 787-1400
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 169 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-895-939-4
;
; Query Match 16.2%; Score 180.5; DB 2; Length 169;
; Best Local Similarity 32.9%; Pred. No. 2.6e-12;
; Matches 53; Conservative 31; Mismatches 54; Indels 23; Gaps 8;
;
QY 46 LEPLTKKETLIAMREFLVNSQ-----DEIAKRYGLSGVGLAAPQINISKRMIAVL 100
Db 7 LHTP---DBRLRKVAKPVVEYNAEIQIIVDDMPETWYA-BEGIGLAATQVDIHQRIIVI- 61
;
QY 101 IPDDGSKSYDMLVNPKIVSHSVQAYLPTG-EGCLSVDDNVAAGLVHRNKITTKAKD 158
Db 62 --DVSNNRDERLVLINPELEKSGE-----TGIEGCLSIPEQRA-LVPAEKVKIRALD 113
;
QY 159 IEGNDIQRLKGYPAIVFOHEIDHLNGVMFYDHIKDHPLQ 199
Db 114 RDGKPFLEADGGLAICIQHEMDHVLGKLFMDYLS---PLK 151
;
;
; RESULT 10
; US-09-188-820-4
; Sequence 4, Application US/09188820
; Patent No. 6190902
; GENERAL INFORMATION:
; APPLICANT: Bogosian, Gregg
; TITLE OF INVENTION: Deformylation of f-Met Peptides in
; Bacterial Expression Systems
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/188,820
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/895,939
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Patterson, Melinda L.
REGISTRATION NUMBER: 33,062
REFERENCE/DOCKET NUMBER: MOPV:006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 787-1400
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-188-820-4

Query Match 16.2%; Score 180.5; DB 4; Length 169;
Best Local Similarity 32.9%; Pred. No. 2.6e-12;
Matches 53; Conservative 31; Mismatches 54; Indels 23; Gaps 8;
QY 46 LELPTKEKETIAMREFLVNSQ-----DEETAKRYGLRSGVGLAAPQINISKRMIAVL 100
Db LHP---DEELKVKAPVEEVNAEIQRIYDDMFETMYA-ESGIGLAATQVDIHQRIIVI- 61
QY 101 IPDDGSKSYDMLNPKIVSHSVQEAFLPTG--EGCLSVDDNVAGLVHRRNKITIKAD 158
Db --DVSENDRERLVNINPLELKSQE-----TGIEEGCLSIPEQRA-LVPRAEKVKIRALD 113
QY 159 IEGNDIQLRLKGYPAIVFQHEIDHLNGVMFVYDHDKDHPLO 199
Db LRDGPFELEADGLLAICIQHEMDHLVGVLFMDYLS---PLK 151

RESULT 11
US-09-134-001C-4828
Sequence 4828, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4828
LENGTH: 163
TYPE: PRP
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4828

Query Match 13.2%; Score 147; DB 4; Length 163;
Best Local Similarity 25.3%; Pred. No. 1.3e-08;
Matches 46; Conservative 41; Mismatches 71; Indels 24; Gaps 7;
QY 25 MLMDKDIIRDGHPTLRQKAAELEPLTKKETLIAMREFLVNSQDEETAKRYGLRSGVG 84
Db VMIVKLVKSTPHILNTIQ-----FVSTYDQKLVLEDDLTLYHEEA-----A 47

QY 85 LAAPQINISKRMIAVLIPDDGSKSYDMLNPKIVSHSVQEAFLPTGEGCLSVDDNVAG 144
Db ISAPQIGVDQSVAILDMBOEGL-----LQINPVVKSQS-QET-VSDLEGSISL-PHYIG 99
QY 145 LVHRHNKITIKAKDIEGNDIQLRLKGYPAIVFQHEIDHLNGVMF---YDHDKDHPLOPH 201
Db 100 EYKRSKMITVQSYDINGNAVELTAYDDIARMLHWDHLNGIQFTKRAHILNEVEZAY 159
QY 202 TD 203
Db 160 FD 161

RESULT 12
US-08-248-021A-2
Sequence 2, Application US/08248021A
Patent No. 5648240
GENERAL INFORMATION:
APPLICANT: Hook, Magnus
APPLICANT: Jonsson, Klas
APPLICANT: Patti, Joseph M.
APPLICANT: Gurusiddappa, Sivashankarappa
TITLE OF INVENTION: MHC II ANALOG FROM STAPHYLOCOCCUS AUREUS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/248,021A
FILING DATE: 24-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: TAMK:155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 689 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-248-021A-2

Query Match 8.0%; Score 89; DB 1; Length 689;
Best Local Similarity 21.2%; Pred. No. 0.35;
Matches 36; Conservative 27; Mismatches 63; Indels 44; Gaps 4;
QY 4 YLNNLFTVIQLKQIKRKVQYMLTMKDII RDGHPTRLRQKAAELEPLTKKETLIAMRE 63
Db FTRNLFVSKDIKIDINVKQSKSKNALNK-----VTNKATKVKFPVTINGFSNLVS-NE 610
QY 64 FL-----VNSODEETAKRYGLRSGVGLAAPQINISKRMIAVLIPDDGSKSYDMLVNP 117
Db FAFLPHKITTNDLNAKLRALRSQGITKHDIGLSERTVYKVFYKDGSSKLEDLKAA-- 668
QY 118 KIVSHSVQEAFLPTGEGCLSVDDNVAGLVHRRNKITIKAKDIEGNDIQLR 167
Db 669 -----KQDSKVFKATDKKKVDIEIK 688

RESULT 13

US-08-469-005A-10
; Sequence 10, Application US/08469005A
; Patent No. 5665874
; GENERAL INFORMATION:
; APPLICANT: KUHAJDA, FRANCIS P.
; APPLICANT: PASTERNAK, GARY A.
; TITLE OF INVENTION: CANCER RELATED ANTIGEN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,005A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/188,426
; FILING DATE: 24-JAN-1994
; APPLICATION NUMBER: 08/096,908
; FILING DATE: 26-JUL-1993
; APPLICATION NUMBER: 07/917,716
; FILING DATE: 24-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 062482-0113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2509 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-469-005A-10

Query Match 7.6%; Score 84; DB 1; Length 2509;
Best Local Similarity 21.6%; Pred. No. 9.3;
Matches 27; Conservative 32; Mismatches 40; Indels 26; Gaps 3;
QY 5 LNNLFTVIQLKQIKRKQVYMLTKDIIRDGHFTLQKAAELEPLTYE-----EKETLIA 60
DB 2168 LNLVLSVREVROTLTKLQELSSKAD-----EASELACPTPKEDGLAQOQTQLN 2216
QY 61 MREFLVNSQDEEIAKRYGLRSVGGLAAPQINISKRMIAVLIPDDGSGKSYDMLVNPXIV 120
DB 2217 LRSLLVNPBGPTLMRLNSVQS-----SERPLFVHPHIEGSTTVFHSIASRLSIP 2265
QY 121 SHSVQ 125
DB 2266 TYGLQ 2270
RESULT 14
US-09-261-907-2
; Sequence 2, Application US/09261907A
; Patent No. 6294364

; GENERAL INFORMATION:
; APPLICANT: ELLIS, CATHERINE
; APPLICANT: LONSDALE, JOHN
; APPLICANT: BERGSMAN, DEK J.
; APPLICANT: MOONEY, JEFFREY L.
; APPLICANT: DEPIERA, MEGAN E.
; APPLICANT: CHAPMAN, CONRAD
; TITLE OF INVENTION: HUMAN FAS
; FILE REFERENCE: GP-70603
; CURRENT APPLICATION NUMBER: US/09/261,907A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2511
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; US-09-261-907-2

Query Match 7.6%; Score 84; DB 4; Length 2511;
Best Local Similarity 21.6%; Pred. No. 9.3;
Matches 27; Conservative 32; Mismatches 40; Indels 26; Gaps 3;
QY 5 LNNLFTVIQLKQIKRKQVYMLTKDIIRDGHFTLQKAAELEPLTYE-----EKETLIA 60
DB 2170 LNLVLSVREVROTLTKLQELSSKAD-----EASELACPTPKEDGLAQOQTQLN 2218
QY 61 MREFLVNSQDEEIAKRYGLRSVGGLAAPQINISKRMIAVLIPDDGSGKSYDMLVNPXIV 120
DB 2219 LRSLLVNPBGPTLMRLNSVQS-----SERPLFVHPHIEGSTTVFHSIASRLSIP 2267
QY 121 SHSVQ 125
DB 2268 TYGLQ 2272

RESULT 15
US-08-557-122A-38
; Sequence 38, Application US/08557122A
; Patent No. 5879664
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten Mailand
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5879664o No. 5879664disk of No. 5879664th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,122A
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3980.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

Search completed: July 11, 2003, 09:42:02
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 09:40:50 ; Search time 51 Seconds
(without alignments)
486.224 Million cell updates/sec

Title: US-09-896-580A-1

Perfect score: 1112
Sequence: 1 MYEYLNLETVIQLKQIKR.....KDHPLQPTDAVEHQHHH 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445759

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07 NEW PUB.pbp.*
2: /cgn2_6/ptodata/1/pubpaa/PCT NEW PUB.pbp.*
3: /cgn2_6/ptodata/1/pubpaa/US06 NEW PUB.pbp.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pbp.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pbp.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pbp.*
7: /cgn2_6/ptodata/1/pubpaa/US08 NEW PUB.pbp.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pbp.*
9: /cgn2_6/ptodata/1/pubpaa/US09 NEW PUB.pbp.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pbp.*
11: /cgn2_6/ptodata/1/pubpaa/US10 NEW PUB.pbp.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pbp.*
13: /cgn2_6/ptodata/1/pubpaa/US60 NEW PUB.pbp.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pbp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	457.5	41.1	203	10	US-09-862-005-2
2	386.5	34.8	155	10	US-09-862-005-4
3	161	14.5	216	9	US-10-156-761-10570
4	160.5	14.4	193	9	US-09-738-626-6508
5	160.5	14.4	193	10	US-09-825-345-2
6	146.5	13.2	224	9	US-10-156-761-14878
7	143	12.9	169	9	US-09-738-626-5271
8	140	12.6	243	9	US-10-149-256-6
9	140	12.6	243	10	US-09-866-099-2
10	139	12.5	186	9	US-10-156-761-14655
11	98.5	8.9	210	9	US-10-149-256-4
12	91	8.2	162	9	US-09-921-144-16
13	91	8.2	163	9	US-09-921-144-14
14	91	8.2	184	9	US-09-921-144-20
15	91	8.2	185	9	US-09-921-144-18
16	87.5	7.9	604	9	US-09-949-293-30
17	86	7.7	522	9	US-09-995-749A-11
18	84.5	7.6	604	9	US-09-949-293-26
19	84	7.6	2509	9	US-10-237-271-1

20	81.5	7.3	544	9	US-09-949-293-29	Sequence 29, Appl
21	81	7.3	896	9	US-10-210-296-5	Sequence 5, Appl
22	80	7.2	163	9	US-10-149-256-2	Sequence 2, Appl
23	80	7.2	389	12	US-10-062-254-300	Sequence 300, App
24	79	7.1	588	9	US-10-156-761-14205	Sequence 14205, A
25	78.5	7.1	600	10	US-09-974-592-12	Sequence 12, Appl
26	78.5	7.1	604	9	US-09-949-293-28	Sequence 28, Appl
27	78.5	7.1	604	9	US-09-953-067A-8	Sequence 8, Appl
28	78	7.0	935	10	US-09-815-242-11358	Sequence 11358, A
29	78	7.0	4349	9	US-10-160-758-15	Sequence 15, Appl
30	77	6.9	468	10	US-09-925-300-1661	Sequence 1661, Ap
31	76	6.8	543	9	US-09-738-626-4484	Sequence 4484, Ap
32	76	6.8	299	9	US-10-225-567A-426	Sequence 426, App
33	75.5	6.8	599	9	US-09-949-293-31	Sequence 31, Appl
34	75.5	6.8	602	9	US-09-953-067A-7	Sequence 40, Appl
35	75.5	6.8	823	9	US-10-177-293-240	Sequence 240, App
36	75.5	6.8	823	9	US-09-976-059-15	Sequence 15, Appl
37	75.5	6.8	4999	9	US-10-172-399-2	Sequence 2, Appl
38	75	6.7	215	9	US-09-738-626-6888	Sequence 6888, Ap
39	75	6.7	1344	9	US-10-205-823-419	Sequence 419, App
40	75	6.7	1979	9	US-10-205-823-419	Sequence 266, App
41	74.5	6.7	350	9	US-10-097-340-266	Sequence 3, Appl
42	74.5	6.7	350	9	US-09-251-225-3	Sequence 341, App
43	74.5	6.7	350	9	US-10-205-823-341	Sequence 966, App
44	74.5	6.7	385	9	US-09-925-299-966	Sequence 4598, Ap
45	74.5	6.7	385	9	US-10-106-698-4598	

ALIGNMENTS

RESULT 1
US-09-862-005-2
; Sequence 2, Application US/09862005
; Patent No. US20020058796A1
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: def1
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/862,005
; FILING DATE: 21-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/991,023
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM50010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 203 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:


```
/ APPLICANT: OZAKI, AKIO
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-125
/ CURRENT APPLICATION NUMBER: US/09/738,626
/ CURRENT FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: JP 99/377484
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: JP 00/159162
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: JP 00/280988
/ PRIOR FILING DATE: 2000-08-03
/ NUMBER OF SEQ ID NOS: 7059
/ SOFTWARE: PatentIn ver. 3.0
/ SEQ ID NO 6508
/ LENGTH: 193
/ TYPE: PRT
/ ORGANISM: Corynebacterium glutamicum
US-09-738-626-6508

Query Match          14.4%; Score 160.5; DB 9; Length 193;
Best Local Similarity 29.7%; Pred. No. 1.9e-07;
Matches 52; Conservative 29; Mismatches 63; Indels 31; Gaps 9;

QY 26 LTKMDIIRGHTLROKAAEELPLTKKEETLIAMREFLVNSQDEBIAKRY---GLRSG 82
Db 1 MTRVRIVHGDPVLHNP-----QLVTE-----VSELQELIADMYETMDVANG 44
QY 83 VGLAAPQINISKRMIAVLIPDDGSGKSYDMLVNPKNIVSHSVQEAAYLPTG-----EGCLS 137
Db 45 VGLAANQIGVSKRIFFVPCDD-EGVMHKGCFINPVLTSIPIET-MPADDGSDSEGCLS 102

QY 138 V--DDNVAGLVHRHNKLTIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYD 190
Db 103 VPGEFPTGRAH-WAKVT--GLNKEGEEVSVEAGFLARCQFQHEVGHLDGFLTYD 154

RESULT 5
US-09-825-345-2
/ Sequence 2, Application US/09825345
/ Patent No. US20020106750A1
/ GENERAL INFORMATION:
/ APPLICANT: FARWICK, MIKE
/ APPLICANT: RUTHWACHER, KLAUS
/ APPLICANT: BREHME, JENNIFER
/ APPLICANT: PEPPERLE, WALTER
/ TITLE OF INVENTION: NEW NUCLEOTIDE SEQUENCES WHICH CODE FOR THE DEF GENE
/ FILE REFERENCE: 21123/280107/MAS
/ CURRENT APPLICATION NUMBER: US/09/825,345
/ CURRENT FILING DATE: 2001-04-04
/ PRIOR APPLICATION NUMBER: DE 100 46 228.6
/ PRIOR FILING DATE: 2000-09-19
/ PRIOR APPLICATION NUMBER: DE 101 13 957.8
/ PRIOR FILING DATE: 2001-03-22
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 193
/ TYPE: PRT
/ ORGANISM: Corynebacterium glutamicum
US-09-825-345-2

Query Match          14.4%; Score 160.5; DB 10; Length 193;
Best Local Similarity 29.7%; Pred. No. 1.9e-07;
Matches 52; Conservative 29; Mismatches 63; Indels 31; Gaps 9;

QY 26 LTKMDIIRGHTLROKAAEELPLTKKEETLIAMREFLVNSQDEBIAKRY---GLRSG 82
Db 1 MTRVRIVHGDPVLHNP-----QLVTE-----VSELQELIADMYETMDVANG 44
QY 83 VGLAAPQINISKRMIAVLIPDDGSGKSYDMLVNPKNIVSHSVQEAAYLPTG-----EGCLS 137
Db 45 VGLAANQIGVSKRIFFVPCDD-EGVMHKGCFINPVLTSIPIET-MPADDGSDSEGCLS 102
```

```
QY 138 V--DDNVAGLVHRHNKLTIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYD 190
Db 103 VPGEFPTGRAH-WAKVT--GLNKEGEEVSVEAGFLARCQFQHEVGHLDGFLTYD 154

RESULT 6
US-10-156-761-14878
/ Sequence 14878, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIRA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 14878
/ LENGTH: 224
/ TYPE: PRT
/ ORGANISM: Streptomyces avermitilis
US-10-156-761-14878

Query Match          13.2%; Score 146.5; DB 9; Length 224;
Best Local Similarity 30.0%; Pred. No. 4.7e-06;
Matches 54; Conservative 26; Mismatches 61; Indels 39; Gaps 9;

QY 31 IIRDGHTLROKAAEELPLTKKEETLIAMREFLVNSQDEBIAKRYGLRSGVGLAAPQI 90
Db 35 IVAAGDPVLRGAE---PYDGLQFGLLA---RFV-----EALRLTWHAAPGVGLAAQV 83
QY 91 NISKRMIAVL-----IPD-----GSGKSYDMLVNPKNIVSHSVQEAAYLPTG-----E 133
Db 84 GVGLR-LAVIEDPAPVPEVGAVRGPQPFVLPN-----SYEAVGSDRAAPFE 133
QY 134 GCLSVDDNVAGLVHRHNKLTIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYD 193
Db 134 GCLSVFQWQA-VVARPARVRLTALDEHGRAVDEEPTGWPARIVOHETDHLDDGMLYLDRAE 192

RESULT 7
US-09-738-626-5271
/ Sequence 5271, Application US/09738626
/ Publication No. US20020197605A1
/ GENERAL INFORMATION:
/ APPLICANT: NAKAGAWA, SATOSHI
/ APPLICANT: MIZOGUCHI, HIROSHI
/ APPLICANT: ANDO, SEIKO
/ APPLICANT: HAYASHI, MIKIO
/ APPLICANT: OCHIALI, KEIKO
/ APPLICANT: YOKOI, HARUHIKO
/ APPLICANT: TATEISHI, NAKO
/ APPLICANT: SENOH, AKIHIRO
/ APPLICANT: IKEDA, MASATO
/ APPLICANT: OZAKI, AKIO
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-125
/ CURRENT APPLICATION NUMBER: US/09/738,626
/ CURRENT FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: JP 99/377484
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: JP 00/159162
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: JP 00/280988
```


Matches 35; Conservative 24; Mismatches 49; Indels 4; Gaps 3;
QY 81 SGVGLAAPQINISKMIIVLPDDGSGKSYDMLVNPVKIVSHVQVAYLPTGEGCLSVDD 140
Db 53 NGVGLAANQIGVPLRVFYVDCPDDEVRHVGH-VVNPRLI--EADGVVLRGPSCCLSLFG 109
QY 141 NVAGLVRHNNKTIKADIEGNDIQLRLKGVPAIVFQHEIDHLNGVMFYDHI 192
Db 110 LEAG-TERIDRAVVEGFTTDEGPVRLGTGWFARCLQHECDHLDGVTVDKV 160
RESULT 11
US-10-149-256-4
; Sequence 4, Application US/10149256
; Publication No. US20030003488A1
; GENERAL INFORMATION:
; APPLICANT: Ramakrishnan, Shyam
; TITLE OF INVENTION: Regulation of Human Mitochondrial
; TITLE OF INVENTION: Deforwylase
; FILE REFERENCE: 004974.00739
; CURRENT APPLICATION NUMBER: US/10/149,256
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: PCT/EP00/12110
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/239,106
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: US 60/169,615
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(210)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-149-256-4

Query Match 8.9%; Score 98.5; DB 9; Length 210;
Best Local Similarity 37.7%; Pred. No. 0.14;
Matches 23; Conservative 8; Mismatches 29; Indels 1; Gaps 1;
QY 133 EGCGLVDDNVAGVHRENNKTIKADIEGNDIQLRLKGVPAIVFQHEIDHLNGVMFYDHI 192
Db 3 EGESVAAFXA-CVPRFOXVQLSGLDENGSGVWQASGWAARIIOHENDHILQGLFIDKM 61
QY 193 D 193
Db 62 D 62
RESULT 12
US-09-921-144-16
; Sequence 16, Application US/09921144
; Publication No. US20030027247A1
; GENERAL INFORMATION:
; APPLICANT: ZHONG, Pingyu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR GENERATING
; TITLE OF INVENTION: CHIMERIC HETEROMULTIMERS
; FILE REFERENCE: 13403.0004.NPUS00
; CURRENT APPLICATION NUMBER: US/09/921,144
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Bluescript vector
US-09-921-144-16

Query Match 8.2%; Score 91; DB 9; Length 162;
Best Local Similarity 25.1%; Pred. No. 0.51;
Matches 42; Conservative 27; Mismatches 56; Indels 42; Gaps 8;
QY 50 LTKEEKETLIAMREPLVNSODEEIAK-RYGLRSGVGLAAPQINISKMIIVLPDDGSK 108
Db 35 LEKENREL-----EKIIAEKEERVSERHQLOSVGGCGAPOFHSKEVIMKKLL----- 82
QY 109 SYDYMIVNPVKIVSHVQVAYLPTGEGCLSVDDNVAGLVHR-HNKITTIKADIEGNDIQLR 167
Db 83 -FAIPLWVP-FYSHSATRS--RGGGTSRLEGLQSENRHLRMKITELDKDLEVTMQLQ 137
QY 168 LKGYPAIVFQHEI-DHLNGVMFYDHIKDHPLOPHTDAVEVHQHHH 213
Db 138 DVGGCAAYPYDVPDYAGG-----HHHHH 162
RESULT 13
US-09-921-144-14
; Sequence 14, Application US/09921144
; Publication No. US20030027247A1
; GENERAL INFORMATION:
; APPLICANT: ZHONG, Pingyu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR GENERATING
; TITLE OF INVENTION: CHIMERIC HETEROMULTIMERS
; FILE REFERENCE: 13403.0004.NPUS00
; CURRENT APPLICATION NUMBER: US/09/921,144
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Bluescript vector
US-09-921-144-14

Query Match 8.2%; Score 91; DB 9; Length 163;
Best Local Similarity 25.1%; Pred. No. 0.51;
Matches 42; Conservative 27; Mismatches 56; Indels 42; Gaps 8;
QY 50 LTKEEKETLIAMREPLVNSODEEIAK-RYGLRSGVGLAAPQINISKMIIVLPDDGSK 108
Db 36 LEKENREL-----EKIIAEKEERVSERHQLOSVGGCGAPOFHSKEVIMKKLL----- 83
QY 109 SYDYMIVNPVKIVSHVQVAYLPTGEGCLSVDDNVAGLVHR-HNKITTIKADIEGNDIQLR 167
Db 84 -FAIPLWVP-FYSHSATRS--RGGGTSRLEGLQSENRHLRMKITELDKDLEVTMQLQ 138
QY 168 LKGYPAIVFQHEI-DHLNGVMFYDHIKDHPLOPHTDAVEVHQHHH 213
Db 139 DVGGCAAYPYDVPDYAGG-----HHHHH 163
RESULT 14
US-09-921-144-20
; Sequence 20, Application US/09921144
; Publication No. US20030027247A1
; GENERAL INFORMATION:
; APPLICANT: ZHONG, Pingyu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR GENERATING
; TITLE OF INVENTION: CHIMERIC HETEROMULTIMERS
; FILE REFERENCE: 13403.0004.NPUS00
; CURRENT APPLICATION NUMBER: US/09/921,144
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Bluescript vector
US-09-921-144-20

```
; FEATURE:
; OTHER INFORMATION: Bluescript vector
US-09-921-144-20

Query Match      8.2%; Score 91; DB 9; Length 184;
Best Local Similarity 25.1%; Pred. No. 0.6; Mismatches 56; Indels 42; Gaps 8;
Matches 42; Conservative 27;

QY 50 LTKKEKETLIAMREFLVNSQDEEIAK-RYGLRSGVGLAAPQINISKRMIAVLIPDDGSGK 108
Db 35 LEKENREL-----EKIIAEKERVSSELRHQLQSVGGCGAPQPHSKVLMKKLL----- 82

QY 109 SYDYMVNPKIVSHSVQEAFLPTGCGCLSVDDNVAGLVHR-HNKITIKAKDIEGNDIQLR 167
Db 83 -FAIPLVVP-FYSHSATRS---RGGGGTSRLEGLQSENHRLMKITELDKOLEEVTNQLQ 137

QY 168 LKGYFAIVFOHEI-DHLNGVMFYDHDKDHPQLPHTDAVEVHQHHH 213
Db 138 DVGGCAAYFYDVPDIAGG-----HHHHH 162

RESULT 15
US-09-921-144-18
; Sequence 18, Application US/09921144
; Publication No. US20030027247A1
; GENERAL INFORMATION:
; APPLICANT: ZHONG, Pingyu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR GENERATING
; TITLE OF INVENTION: CHIMERIC HETEROMULTIMERS
; FILE REFERENCE: 13403.0004.NPUS00
; CURRENT APPLICATION NUMBER: US/09/921,144
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Bluescript vector
US-09-921-144-18

Query Match      8.2%; Score 91; DB 9; Length 185;
Best Local Similarity 25.1%; Pred. No. 0.61; Mismatches 56; Indels 42; Gaps 8;
Matches 42; Conservative 27;

QY 50 LTKKEKETLIAMREFLVNSQDEEIAK-RYGLRSGVGLAAPQINISKRMIAVLIPDDGSGK 108
Db 36 LEKENREL-----EKIIAEKERVSSELRHQLQSVGGCGAPQPHSKVLMKKLL----- 83

QY 109 SYDYMVNPKIVSHSVQEAFLPTGCGCLSVDDNVAGLVHR-HNKITIKAKDIEGNDIQLR 167
Db 84 -FAIPLVVP-FYSHSATRS---RGGGGTSRLEGLQSENHRLMKITELDKOLEEVTNQLQ 138

QY 168 LKGYFAIVFOHEI-DHLNGVMFYDHDKDHPQLPHTDAVEVHQHHH 213
Db 139 DVGGCAAYFYDVPDIAGG-----HHHHH 163
```

Search completed: July 11, 2003, 09:48:45
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 09:33:30 ; Search time 40 Seconds
(without alignments)
511.916 Million cell updates/sec

Title: US-09-896-580A-1

Perfect score: 1112

Sequence: 1 MYEVLNNLFTVIQKIKR.....KDHPLQPHDTDAVEHQHHH 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 73:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	938	84.4	183	2 A89879	hypothetical prote
2	575.5	51.8	184	1 D69862	formylmethionine d
3	536.5	48.2	183	2 AB1563	formylmethionine d
4	533.5	48.0	183	2 AC1206	formylmethionine d
5	506.5	45.5	182	2 B83982	formylmethionine d
6	462.5	41.6	203	2 B98035	formylmethionine d
7	461.5	41.5	203	2 B95169	polypeptide deform
8	436	39.2	211	2 B86594	polypeptide deform
9	273.5	24.6	198	2 F82886	polypeptide deform
10	263	23.7	198	2 A90598	hypothetical prote
11	230	20.7	216	2 S73913	polypeptide deform
12	226.5	20.4	224	2 H97709	formylmethionine d
13	222.5	20.0	226	1 G64211	formylmethionine d
14	218.5	19.6	168	2 C82494	polypeptide deform
15	217.5	19.6	179	2 AI2056	polypeptide deform
16	215.5	19.4	191	2 A55228	fms protein homolo
17	211.5	19.0	169	1 D64082	N-formylmethionyla
18	207	18.6	181	2 C81680	polypeptide deform
19	206	18.5	273	2 T48639	hypothetical prote
20	205.5	18.5	232	2 H75274	polypeptide deform
21	203	18.3	181	2 B71526	probable polypepti
22	198	17.8	164	2 C72224	polypeptide deform
23	193.5	17.4	170	2 F97403	polypeptide deform
24	193.5	17.4	170	2 AF2621	polypeptide deform
25	190	17.1	173	2 B84987	formylmethionine d
26	189	17.0	150	2 A97205	N-formylmethionyl-
27	184.5	16.6	160	1 F69613	polypeptide deform
28	184.5	16.6	169	2 AB1010	formylmethionine d
29	184	16.5	259	2 B86288	hypothetical prote

30	182.5	16.4	169	2 C70352	polypeptide deform
31	181.5	16.3	187	2 AF3542	formylmethionine d
32	180.5	16.2	169	2 S23107	N-formylmethionyla
33	180.5	16.2	169	2 H91147	peptide deformylas
34	180.5	16.2	169	2 D85993	polypeptide deform
35	180.5	16.2	187	1 S77378	polypeptide deform
36	178.5	16.1	150	2 E97112	N-formylmethionyl-
37	177.5	16.0	175	2 F81437	formylmethionine d
38	176	15.8	170	2 H82746	polypeptide deform
39	176	15.8	170	2 AC0030	formylmethionine d
40	175	15.7	168	2 H83643	polypeptide deform
41	175	15.7	186	2 H86623	polypeptide deform
42	175	15.7	186	2 E72000	polypeptide deform
43	174.5	15.7	187	2 AH2190	polypeptide deform
44	169.5	15.2	169	2 A82373	polypeptide deform
45	169.5	15.2	173	2 G87282	polypeptide deform

ALIGNMENTS

RESULT 1

A89879

hypothetical protein pdf1 [imported] - Staphylococcus aureus (strain N315)

C/Species: Staphylococcus aureus

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 31-Dec-2001

C/Accession: A89879

R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogi

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A/Reference number: A89758; MUID:21311952; PMID:11418146

A/Accession: A89879

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-183 <KUR>

A/Cross-references: GB:BA000018; PID:gl3700892; PIDN:BA842188.1; GSPDB:GN00149

A/Experimental source: strain N315

C/Genetics:

A/Gene: pdf1

C/Superfamily: polypeptide deformylase

Query Match 84.4%; Score 938; DB 2; Length 183;
Best Local Similarity 98.9%; Pred. No. 1.9e-68;
Matches 181; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 25 MLTMDIIRDGHPTLRQKAAELPLTKKEKETLIAMFEFLVNSQDEETAKRYGLRSGVG 84

Db 1 MLTMDIIRDGHPTLRQKAAELPLTKKEKETLIAMFEFLVNSQDEETAKRYGLRSGVG 60

QY 85 LAAPQINISKMIAMVLPDDGSGKSYDYMVNPKIVSHSVQEAFLPTGEGCLSDVDNVAG 144

Db 61 LAAPQINISKMIAMVLPDDGSGKSYDYMVNPKIVSHSVQEAFLPTGEGCLSDVDNVAG 120

QY 145 LVHRNKKTIKADIEGNDIQFLKGPFAIVFOEHDLNGWFWFYDHDKDHPLOPHDTA 204

Db 121 LVHRNKKTIKADIEGNDIQFLKGPFAIVFOEHDLNGWFWFYDHDKDHPLOPHDTA 180

QY 205 VEV 207

Db 181 VEV 183

RESULT 2

D69862

formylmethionine deformylase homolog ykrB - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C/Accession: D69862

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

[illegible]

RESULT 8
H86694
polypeptide deformylase [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)
C:Species: *Lactococcus lactis* subsp. *lactis*
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: H86694
R:Belotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
A:Reference number: H86625; MUID:21235186; PMID:11337471

124 Q7 Q0EAVLPTGEGCLSVDDNVAGIVHRENKIITIKAKOTEGNDIQLRLKGPVATVFOHEIDHL 183
8 / NINGKGGVGHUATQGLJMKANLWUJLTFNEFEVDGNSLFFASALNTEKELWAKVWVSD 186
127 Q0EAAVPTGEGCLSVDRVPVGVVRHRAVTVTYNKEGEGKKIKLKPALCVQHEIDHT 186

Dd 57 VGVAAVQYGLKKNFYVCVPNDSLTFORDSSQEVRDEKYLFPRDVINPENVIKSDMVA 116

Qy 129 LPTGEGCLSVDD---NAVLAGVHRHNKITIKAKD-IBGNDIQLRLKGYPALVFQHEDHLN 184
 ::| | | | | : : : :
Db 117 ISQGECGLSVDSESWFNQBLAVRRHWEIWKVGYSYFQKKEMTWHWKGYYAIVFOHELDHLN 176
 ::| | | | | : : : :

Qy 185 GWMFPYDHID 193
 || :: || | |
Db 177 GMFLFDRIID 185

RESULT 11

S73913

polypeptide deformylase def - Mycoplasma pneumoniae (strain ATCC 29342)

N;Alternate names: hypothetical protein K04_orf216

C;Species: Rickettsia conorii

A;Variety: ATCC 29342

C;Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999

R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae

A;Reference number: S73327; PMID:97105885; PMID:8948633

A;Accession: S73913

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-216 <HM>

A;Cross-references: EMBL:AEOO0057; GB:U00089; NID:g1674279; PIDN:AB96235.1; PID:g16742

A;Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1996

C;Genetics:

A;Gene: def

C;Superfamily: polypeptide deformylase

Query Match 20.7%; Score 230; DB 2; Length 216;
Best Local Similarity 30.4%; Pred.No. 2.5e-11;
Matches 59; Conservative 44; Mismatches 83; Indels 8; Gaps 5;

Qy 7 NLFTVIQT-LKQIRKRVQMILTMKDITRDGHPTLRCKAAAELEPLTKKEKETLIAMREFL 65
 ::| | | | | : : : : : : :
Db 11 SIFFILLIPQI---NMELLTPKWVLVDVKINEPTKFVPQDFLDQSALDCIAKWAYV 67

Qy 66 VNDSOEAIKRXYGLRSYGVLGAAPQINTSKMIATVLPDPGSOKSYDYMLNPXKVISHVVQ 125
 ::| | | | | : : : : : : :
Db 68 DASTNGPD-AEKTYIGPIGICIAANGIQYWKQPFIHLM--GGVEHKCLLINPKLIINSAN 124

Qy 126 EAYLPTGEGLSDVNAGLVHRHNKITIKAKD-IBGNDIQLRLKGYPALVFQHEDHLN 184
 ::| | | | | : : : : : : :
Db 125 KSFKLSGEGCLSVPKMHQGYVIREHWITTTGFPMLOQQELITTATGLFGMCLOHQEFDLHQ 184
 :|| | | | | | :

Qy 185 GWMFYDHDKDHP 198
 || :: || | |
Db 185 GRFYTHRINFNL 198

RESULT 12

H97709

formylmethionine deformylase (BC 3.5.1.3.) - Rickettsia conorii (strains Malish 7)

C;Species: Rickettsia conorii

C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001

C;Accession: H97709

F; Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; I Science 293, 2093-2098, 2001

A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A;Reference number: A97700; PMID:21442074; PMID:11557893

A;Accession: H97709

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-224 <UR>

A;Cross-references: GB:AE006914; PIDN:AAJ02618.1; PID:g15619118; GSDB:GN00173

C;Genetics:

A;Gene: def3

C;Keywords: hydrolase

Db 110 GLVPRHOAIEVEYTDRYGNLQKQTLTDFVARIFQHEFDHLDGVLFIDRVESN 161

Search completed: July 11, 2003, 09:41:27
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 09:26:10 ; Search time 23 Seconds
(without alignments)
384.107 Million cell updates/sec

Title: US-09-896-580A-1

Perfect score: 1112
Sequence: 1 MYEYLNLFVTLQKQIKR.....KDHLPQHTDAVEVHQHHH 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	575.5	51.8	184	1	DEF2_BACSU
2	540.5	48.6	184	1	DEF2_BACST
3	506.5	45.5	182	1	DEF_BACHD
4	461.5	41.5	203	1	DEF_STRPN
5	426	38.3	196	1	DEF_LACLA
6	230	20.7	216	1	DEF_MYCPN
7	227	20.4	170	1	DEF_PASMO
8	222.5	20.0	216	1	DEF_MYCGE
9	215.5	19.4	192	1	DEF_THETH
10	211.5	19.0	169	1	DEF_HAFIN
11	208	18.7	273	1	DEF_ARATH
12	207	18.6	181	1	DEF_CHLMU
13	203	18.3	181	1	DEF_CHLTR
14	198	17.8	164	1	DEF_THEMA
15	190	17.1	173	1	DEF_BUCAI
16	188	16.9	277	1	DEF_LYCES
17	187	16.8	279	1	DEF_LYCES
18	184.5	16.6	160	1	DEF1_BACSU
19	184	16.5	259	1	DEF1_ARATH
20	182.5	16.4	169	1	DEF_AQUAE
21	180.5	16.2	168	1	DEF_ECOLI
22	180.5	16.2	187	1	DEF_SYNY3
23	178.5	16.1	150	1	DEF_CLOAB
24	177.5	16.0	175	1	DEF_CAMJE
25	176.5	15.9	187	1	DEF_FREDI
26	176	15.8	170	1	DEF_XYLFA
27	175	15.7	168	1	DEF_PSEAE
28	175	15.7	186	1	DEF_CHLPN
29	169.5	15.2	169	1	DEF_VIRCH
30	163	14.7	167	1	DEF_NEICM
31	155	13.9	175	1	DEF_RICET
32	151.5	13.6	174	1	DEF_HELPF
33	148.5	13.4	162	1	DEF_TREPA

34 147.5 13.3 172 1 DEF_BORBU
35 147.5 13.3 174 1 DEF_HELPJ
36 146 13.1 175 1 DEF_RICPR
37 140 12.6 243 1 DEFM_HUMAN
38 137.5 12.4 208 1 DEF_STRCO
39 106 9.5 136 1 DEF_CLOBE
40 105.5 9.5 197 1 DEF_MYCLE
41 105.5 9.5 197 1 DEF_MYCTU
42 87.5 7.9 604 1 PGH2_HORSE
43 87 7.8 880 1 SYV_BACHD
44 86.5 7.8 364 1 APAR_BACST
45 85.5 7.7 445 1 ASSY_PASMU

ALIGNMENTS

RESULT 1
DEF2_BACSU
ID DEF2_BACSU STANDARD; PRT; 184 AA.
AC Q45495;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide deformylase 2 (EC 3.5.1.88) (PDF 2) (Polypeptide deformylase 2).
DE 2).
GN DEF2.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Caldwell R.M., Ferrari B.;
RT "Sequence analysis of the mobA-ampS region of the Bacillus subtilis chromosome."
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Comerston I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly B.J., Grandi G., Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapilus A., Lardinols S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Roone D., O'Reilly K., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y., Sato T., Scallan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat K., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.P., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis."
RL Nature 390:249-256(1997).
RN [3]

CHARACTERIZATION.
MEDLINE=2132705; PubMed=11429456;
RA Haas M., Beyer D., Gahlmann R., Freiberg C.;
RT "YkrB is the main peptide deformylase in Bacillus subtilis, a
RL eubacterium containing two functional peptide deformylases.";
CC Microbiology 147:1783-1791 (2001).
CC -1- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions.
CC -1- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -1- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -1- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF012285; AAC24930.1; -
CC EMBL: Z99111; CAB13329.1; -
CC HSSP: P27251; 2DEF.
CC Subtilast; BGL1815; defB.
CC InterPro: IPR000181; Pep deformylase.
CC Pfam: PF01327; Pep deformylase; 1.
CC ProDom: PD003844; Pep deformylase; 1.
CC TIGRFAMs: TIGR00079; Pept deformyl; 1.
CC Protein biosynthesis; Hydrolase; Iron; Complete proteome.
CC METAL 110 110 IRON (BY SIMILARITY).
CC FT METAL 153 153 IRON (BY SIMILARITY).
CC FT ACT SITE 154 154 BY SIMILARITY.
CC FT METAL 157 157 IRON (BY SIMILARITY).
CC SQ SEQUENCE 184 AA; 20655 MW; 8641BF1932666C38 CRC64;

Query Match 51.8%; Score 575.5; DB 1; Length 184;
Best Local Similarity 59.5%; Pred. No. 1.7e-40;
Matches 110; Conservative 33; Mismatches 41; Indels 1; Gaps 1;

QY 25 MVTWKDIIRDCPTLRQKALELPKTEKSTLIAMREFLVNSODEIAKRYGLRSVG 84
Db 1 MITMENIVRQGHFALRETPVELPTDAERQQLADMFVRSQNFELAKYKIRFGVG 60

QY 85 LAAPQINISKRMIAVLIPDGGSKSYDMLVNPKEVSHSVQEAAYLPTGEGCLSYDDNVAG 144
Db 61 LAAPQINISKRMIAV-HAEDASGLYSVALFNPKEVSHSVQEAAYLPTGEGCLSYDDNVAG 119

QY 145 LVHRENKTIKADIEGNDIQRLKGYPAIVFOHEIDLHNGVMFYDHDKHPLQHTDA 204
Db 120 YVFRARIVRGTTLEGNDIRLKGFFAIVFOHEIDLHNGVMFYDHDKENPKFEPENA 179

QY 205 VEVHQ 209
Db 180 IAIER 184

RESULT 2
DEF2_BACST
ID DEF2_BACST STANDARD; PRT; 184 AA.
AC O31410;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide deformylase 2 (EC 3.5.1.88) (PDF 2) (Polypeptide deformylase
2).
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]

SEQUENCE FROM N.A., AND CHARACTERIZATION.
STRAIN=ATCC 1518;
RX MEDLINE=97272005; PubMed=9126850;
RA Meinel T., Lazennec C., Villoing S., Blanquet S.;
RT "Structure-function relationships within the peptide deformylase
RT family. Evidence for a conserved architecture of the active site
RT involving three conserved motifs and a metal ion.";
RL J. Mol. Biol. 267:749-761 (1997).
CC -1- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions.
CC -1- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -1- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -1- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y10549; CAAT1581.1; -
CC HSSP: P27251; IDEF.
CC InterPro: IPR000181; Pep deformylase.
CC Pfam: PF01327; Pep deformylase; 1.
CC ProDom: PD003844; Pep deformylase; 1.
CC TIGRFAMs: TIGR00079; Pept deformyl; 1.
CC Protein biosynthesis; Hydrolase; Iron.
CC METAL 110 110 IRON (BY SIMILARITY).
CC FT METAL 153 153 IRON (BY SIMILARITY).
CC FT ACT SITE 154 154 BY SIMILARITY.
CC FT METAL 157 157 IRON (BY SIMILARITY).
CC SQ SEQUENCE 184 AA; 20382 MW; 9CD85DEB53632FA0 CRC64;

Query Match 48.6%; Score 540.5; DB 1; Length 184;
Best Local Similarity 56.8%; Pred. No. 1.3e-37;
Matches 104; Conservative 33; Mismatches 45; Indels 1; Gaps 1;

QY 25 MVTWKDIIRDCPTLRQKALELPKTEKSTLIAMREFLVNSODEIAKRYGLRSVG 84
Db 1 MITWKDIIRDCPTLRQKALELPKTEKSTLIAMREFLVNSODEIAKRYGLRSVG 60

QY 85 LAAPQINISKRMIAVLIPDGGSKSYDMLVNPKEVSHSVQEAAYLPTGEGCLSYDDNVAG 144
Db 61 LAAPQINISKRMIAV-HAEDASGLYSVALFNPKEVSHSVQEAAYLPTGEGCLSYDDNVAG 119

QY 145 LVHRENKTIKADIEGNDIQRLKGYPAIVFOHEIDLHNGVMFYDHDKHPLQHTDA 204
Db 120 YVFRARIVRGTTLEGNDIRLKGFFAIVFOHEIDLHNGVMFYDHDKENPKFEPENA 179

QY 205 VEV 207
Db 180 IPI 182

RESULT 3
DEF2_BACHD
ID DEF2_BACHD STANDARD; PRT; 182 AA.
AC Q9K319;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR B2658.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]

```

RP SEQUENCE FROM N.A.
PC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (By similarity).
CC -1- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -1- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -1- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC EMBL; AP001516; BAB06377.1; -
CC HSP; P27251; 2DEF.
DR InterPro; IPR000181; Pep deformylase.
DR Pfam; PF01327; Pep deformylase; 1.
DR ProDom; PD003844; Pep deformylase; 1.
DR TIGRfams; TIGR00079; Pept deformyl; 1.
KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT METAL 110 110 IRON (BY SIMILARITY).
FT METAL 153 153 IRON (BY SIMILARITY).
FT ACT SITE 154 154 BY SIMILARITY.
FT METAL 157 157 IRON (BY SIMILARITY).
SQ SEQUENCE 182 AA; 20599 MW; 665D39B5DE6153F CRC64;

Query Match 45.5%; Score 506.5; DB 1; Length 182;
Best Local Similarity 58.0%; Pred. No. 7.9e-35;
Matches 98; Conservative 31; Mismatches 39; Indels 1; Gaps 1;

QY 25 MLTKMDITRDGHTPLRQKAELELPLTKKEETLIAMREFLVNSQDELIARVGLRSGV 84
Dd 1 MLTKMDIVREGNVLREAVKVPVPLSDKQTKAKEMLEFLINSQPEIAEKYSLRPGV 60

QY 85 LAAPQINISKRMIAVLIPDGGSGKSYDYMVNPKIVSHVQEAIVLPTGGCLSVDDNVAG 144
Dd 61 LAAPQIGLSKQMIAMVHTDENE-KEYSLVLEFNPKIISVEMTLEGEGCLSVDRVQG 119

QY 145 LVHRNKTIKADIEGNDIQLRLKGYPAIVFQHEIDHNGVWFYDHD 193
Dd 120 IVPFHARITVKAINENNEVRVLKRGFPAPVAFQHEIDHNGIMFYDRIE 168

RESULT 4
DEF_STRPN STANDARD; PRT; 203 AA.
AC Q9F2F0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR SP1456.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13113;
FN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21156952; PubMed=11257016;

```

```

RA Apfel C.M., Locher H., Evers S., Takacs B., Hubschwerlen C.,
RA Pirson W., Page M.G., Keck W.;
RT "Peptide deformylase as an antibacterial drug target: target
RT validation and resistance development.";
RL Antimicrob. Agents Chemother. 45:1058-1064(2001).
RN [2]
RP SEQUENCE FROM N.A.
PC STRAIN=TIG84;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Urkin A.S., Gwim M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Mayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzaple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L., E.K.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
CC -1- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (By similarity).
CC -1- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -1- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -1- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC EMBL; AJ278785; CAC15392.1; -
CC EMBL; AS007442; AAK75550.1; -
CC HSP; P27251; 2DEF.
DR TIGR; SP1456; -
DR InterPro; IPR000181; Pep deformylase.
DR Pfam; PF01327; Pep deformylase; 1.
DR ProDom; PD003844; Pep deformylase; 1.
DR TIGRfams; TIGR00079; Pept deformyl; 1.
KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT METAL 130 130 IRON (BY SIMILARITY).
FT METAL 173 173 IRON (BY SIMILARITY).
FT ACT SITE 174 174 BY SIMILARITY.
FT METAL 177 177 IRON (BY SIMILARITY).
SQ SEQUENCE 203 AA; 22676 MW; 192F91B42375F8CD CRC64;

Query Match 41.5%; Score 461.5; DB 1; Length 203;
Best Local Similarity 50.8%; Pred. No. 4.5e-31;
Matches 95; Conservative 37; Mismatches 48; Indels 7; Gaps 2;

QY 18 KIRKQVYMLTWKDIIRDCHPTLRQKAELELPLTKKEETLIAMREFLVNSQDELIARV 77
Dd 6 RITKKAHLIDNDIIRGNPTLRATAEVTPFLSDQBIILGEGKQMFQKHSQDPVWAEKN 65

QY 78 GLRSGVGLAAPQINISKRMIAVLIPD-----DGGKSYDY--MLVNPKIVSHVQEAIVL 130
Dd 66 GLRSGVGLAAPQIDISKRIIVLVNIVVEGETQEAIDLEALMYNPKIVSHVQDAALG 125

QY 131 TEGGCLSVDDNVAGLVHRNKTIKADIEGNDIQLRLKGYPAIVFQHEIDHNGVWFYD 190
Dd 126 EGGGCLSVDRNVPYVVRARVTVDFDKGKRIKLGKNSIVVQHEIDHNGIMFYD 185

QY 191 HIKDHP 197
Dd 186 RINEKDP 192

```

Db 61 LAANQLGLKKVIAVLIPNEPEVDEDEGNIPPEKAYKREIMYNKAVVSHVQDAAVEGG 120

QY 133 EGSLSVDNVLGVHHRNKITIKAKDIEGNDIOQLRKGYPAIVFQHEIDHLNGVWYFDHI 192

Db 121 EGSLSVDNVLGVHHRNKITIKAKDIEGNDIOQLRKGYPAIVFQHEIDHLNGVWYFDHI 180

QY 193 DKDHP 197

Db 181 NNNDP 185

RESULT 6

DEF MYCEN ID DEF MYCEN STANDARD; PRT; 216 AA.

AC P75527;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).

GN DEF OR MPN246 OR MF587.

OS Mycoplasma pneumoniae.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2104;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ARCC 29342 / M129;

EX MEDLINE=97105885; PubMed=8948633;

RA Hammeleirich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,

RA Herrmann R.;

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma

RT pneumoniae";

RL Nucleic Acids Res. 24:4420-4449 (1996).

CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of

CC newly synthesized proteins. Requires at least a dipeptide for an

CC efficient rate of reaction. N-terminal L-methionine is a

CC prerequisite for activity but the enzyme has broad specificity at

CC other positions (By similarity).

CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +

CC methionyl peptide.

CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).

CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AE000057; BAB96235.1; -.

DR HSPF; P27251; 2DEF.

DR InterPro; IPR000181; Pep_deformylase.

DR Pfam; PF01327; Pep_deformylase; 1.

DR ProDom; PD003844; Pep_deformylase; 1.

DR TIGRFAMs; TIGR00079; Pept_deformyl; 1.

KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.

FT METAL 134 134 IRON (BY SIMILARITY).

FT METAL 178 178 IRON (BY SIMILARITY).

FT ACT_SITE 179 179 BY SIMILARITY.

FT METAL 182 182 IRON (BY SIMILARITY).

SQ SEQUENCE 216 AA; 24592 MW; 361F43404B505DF CRC64;

Query Match

Best Local Similarity 20.7%; Score 230; DB 1; Length 216;

Matches 59; Conservative 44; Mismatches 83; Indels 8; Gaps 5;

QY 7 NLEPTVQ-LKQIKTKRVQVYMLTKDIIRGHPTFRQKAAELFLTKRKETLIARBEFL 65

Db 11 SIFRIILIPQI---NMELPTKAWLVLDVKEINPTKPVQFLDQASIDCIANWAV 67

QY 66 VNSQDEETAKYRGVSGVLAAPQINISIKMTAVLI PDDGSGKSYDMLVNPKITVSHSVQ 125

RESULT 5

DEF LACTA

AC Q48661; Q9C108; STANDARD; PRT; 196 AA.

DT 30-MAY-2000 (Rel. 39, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).

GN DEF OR LI0560.

OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

OX NCBI_TaxID=1360;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LM0230;

EX MEDLINE=95291425; PubMed=7773380;

RA Cancilla M.R., Hillier A.J., Davidson B.E.;

RT "Lactococcus lactis glyceroldehyde-3-phosphate dehydrogenase gene,

RT gap: further evidence for strongly biased codon usage in glycolytic

RT pathway genes";

RL Microbiology 141:1027-1036 (1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Ili1403;

EX MEDLINE=21235186; PubMed=11337471;

RA Boletín A., Wincker P., Mauger S., Jaillon O., Malarne K.,

RA Weissenbach J., Ehrlich S.D., Sorokin A.;

RT "The complete genome sequence of the lactic acid bacterium Lactococcus

RT lactis ssp. lactis Ili1403";

RL Genome Res. 11:731-753 (2001).

CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of

CC newly synthesized proteins. Requires at least a dipeptide for an

CC efficient rate of reaction. N-terminal L-methionine is a

CC prerequisite for activity but the enzyme has broad specificity at

CC other positions (By similarity).

CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +

CC methionyl peptide.

CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).

CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; L36907; AAC41454.1; ALT_INIT.

DR EMBL; AE006230; AAK04658.1; ALT_INIT.

DR InterPro; IPR000181; Pep_deformylase.

DR Pfam; PF01327; Pep_deformylase; 1.

DR ProDom; PD003844; Pep_deformylase; 1.

DR TIGRFAMs; TIGR00079; Pept_deformyl; 1.

KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.

FT METAL 123 123 IRON (BY SIMILARITY).

FT METAL 166 166 IRON (BY SIMILARITY).

FT METAL 167 167 BY SIMILARITY.

FT ACT_SITE 170 170 IRON (BY SIMILARITY).

FT METAL 170 170 IRON (BY SIMILARITY).

FT CONFLICT 86 86 D -> E (IN REF. 1).

SQ SEQUENCE 196 AA; 22047 MW; C5914BB3A92BF243 CRC64;

Query Match

Best Local Similarity 38.3%; Score 426; DB 1; Length 196;

Matches 91; Conservative 36; Mismatches 46; Indels 12; Gaps 3;

QY 25 MLTMKDIIRGHPTFRQKAAELFLTKRKETLIARBEFLVNSQDEETAKYRGVSGV 84

Db 1 MISMDIIRRGVPTLRVANDVTLPSDEILIGKMLQFLNSQDPVAEKGLRGV 60

QY 85 LAAPQINISIKMTAVLI PDDGSGKSYDMLVNPKITVSHSVQAVLPTG 132


```

Db      68 DASNGD-AEKYGIPIGIGIAANQIYKQMFYIHLMD--GGVHEKCLLINPKIINLSAN 124
QY      126 EAYLPTGCGCLSVDDNAGVLRHKNKTIKAKD-IEGNDIQRLKGYPAIVFQHEIDHLN 184
Db      125 KSFLKSGEGLSVPKQKHQYVIRHEWITITGFDWLQOKEITITATGLFGMCLQEPHLQ 184
QY      185 GWFYDHDIXOHEL 198
Db      185 GRFYHRIINPL 198

RESULT 7
DEF_PASMU
ID DEF_PASMU STANDARD; PRT; 170 AA.
AC P57348;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR PM1559.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Pausian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (by similarity).
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (by similarity).
CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB006193; AAK03643.1; -.
DR HSSP; P27251; 2DEF.
DR InterPro; IPR000181; Pep deformylase.
DR Pfam; PF01327; Pep deformylase; 1.
DR ProDom; PD003844; Pep deformylase; 1.
DR TIGRPFAMs; TIGR000079; Pept deformyl; 1.
KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT METAL 91 91 IRON (BY SIMILARITY).
FT METAL 133 133 IRON (BY SIMILARITY).
FT ACT_SITE 134 134 BY SIMILARITY.
FT METAL 137 137 IRON (BY SIMILARITY).
SQ SEQUENCE 170 AA; 19385 MW; 2A65407534AD5A94 CRC64;

Query Match 20.4%; Score 237; DB 1; Length 170;
Best Local Similarity 37.7%; Pred. No. 6 3e-12;
Matches 61; Conservative 26; Mismatches 57; Indels 18; Gaps 6;

QY 44 AELEPLTKKEKTELIAAREFLVNSQDEEIA----KEYGLRSGVGLAAPQINISKRMIAY 99
Db 2 ARLANLVYPDELKIIAKPWEVNDREIIVDNNFMFTWLEEGIGLAAPQVNIHQRIITI 61

QY 100 LIPDDGSKSDYMLVNPXIVSHSVQEAFLPTG--EGLSVDDNAGVLRHKNKTIKAK 157
Db 62 ---DVECTKENQYVLINPEIIDSQGE-----TGIEGCLSL-PGRFGFVPRKEKVTIKAL 112

```

```

QY 158 DIEGNDIQRLKGYPAIVFQHEIDHLNWMFYDHDIXOHELQ 199
Db 113 DRHGEYTLISAGLLAICQHEIDHLNGIVFADYLS---PLK 151

RESULT 8
DEF_MYCGE
ID DEF_MYCGE STANDARD; PRT; 216 AA.
AC P47352;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR MG106.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AUCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Uutterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-P., Dougherty B.A., Bost K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL Science 270:397-403(1995).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (by similarity).
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (by similarity).
CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U39691; AAC71324.1; ALT_INIT.
DR HSSP; P27251; 2DEF.
DR TIGR; MG106; -.
DR InterPro; IPR000181; Pep deformylase.
DR Pfam; PF01327; Pep deformylase; 1.
DR ProDom; PD003844; Pep deformylase; 1.
DR TIGRPFAMs; TIGR000079; Pept deformyl; 1.
KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT METAL 134 134 IRON (BY SIMILARITY).
FT METAL 178 178 IRON (BY SIMILARITY).
FT ACT_SITE 179 179 BY SIMILARITY.
FT METAL 182 182 IRON (BY SIMILARITY).
SQ SEQUENCE 216 AA; 25171 MW; 52D22E8FA58BEFC5 CRC64;

Query Match 20.0%; Score 222.5; DB 1; Length 216;
Best Local Similarity 31.5%; Pred. No. 2e-11;
Matches 64; Conservative 35; Mismatches 95; Indels 9; Gaps 5;

QY 1 NYEYNNLFTVIQ--LKOIKIRKQVYMLTKOIRDCHPTLFQKAAELEPLTKKEKTEL 58
Db 1 MTKILPWLFTSVIRIITLTLFLSMTFQPTKTWLVPDDNALINKPTEAVNPFI-DEQIETC 59

QY 59 IAMEFVLNSQDEETAKYGRVRSVGLAAPQINISGRMIAYLIPDDGSKSDYMLVNPXK 118
Db 60 IKKMIAYDASVDGKAOEYDIPIGIGIAANQIYKQMFYIHLMD--LNKKEKCLLINPK 117

```



```

RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.B., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RL genome sequence of Thermotoga maritima.";
RL Nature 393:323-329(1999).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (By similarity).
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Y10306; CAAY1356.1; -.
DR EMBL; AE001808; RAD36728.1; -.
DR HSPSP; P27251; 2DEF.
DR TIGR; TMI661; -.
DR InterPro; IPR000181; Pep deformylase.
DR Pfam; PF01327; Pep deformylase; 1.
DR ProDom; PD003844; Pep deformylase; 1.
DR TIGRFAMs; TIGR00079; Pept deformyl; 1.
KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT METAL 87 87 IRON (BY SIMILARITY).
FT METAL 129 129 IRON (BY SIMILARITY).
FT ACT_SITE 130 130 BY SIMILARITY.
FT METAL 133 133 IRON (BY SIMILARITY).
SQ SEQUENCE 164 AA; 19024 MW; 9FE94A206DE50842 CRC64;

Query Match 17.8%; Score 198; DB 1; Length 164;
Best Local Similarity 34.2%; Pred. No. 1.4e-09;
Matches 55; Conservative 28; Mismatches 50; Indels 28; Gaps 7;

QY 35 GHTPLRQKAAELPLTKKEE---KETLIAMRFLVNSQDEIAKRYGLRSGVGLAAPQIN 91
DB 8 GDPVLKRAK---PVYFEDENLKKTIEMTWIH-----YDGVGLAAPQVG 51

QY 92 ISKRMIAVLIPDDGSGKSYDMLVNPKIYSHVQRAYLFTGEGCLSVDDNVAGLVHRHNK 151
DB 52 ISQRF---FVMDVNGG---PVAVINPEIL--EIDPETEVAEGCLSPFE-IFVETIERSKR 102

QY 152 ITIKAKDIEGNDIQRLKGYPAIVFQHEIDHLNGVMFYDHI 192
DB 103 IKVYQNTREYVEELEGYAARVFQHEFDHLNGVLIIDRI 143

RESULT 15
DEF_BUCAL
ID DEF_BUCAL STANDARD; PRT; 173 AA.
AC P57563;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR BU496
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;

```

```

EX MEDLINE-20445173; PubMed-10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RL Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (By similarity).
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AP001119; BAB13189.1; -.
DR HSPSP; P27251; 1DEF.
DR InterPro; IPR000181; Pep deformylase.
DR Pfam; PF01327; Pep deformylase; 1.
DR ProDom; PD003844; Pep deformylase; 1.
DR TIGRFAMs; TIGR00079; Pept deformyl; 1.
KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT METAL 91 91 IRON (BY SIMILARITY).
FT METAL 133 133 IRON (BY SIMILARITY).
FT ACT_SITE 134 134 BY SIMILARITY.
FT METAL 137 137 IRON (BY SIMILARITY).
SQ SEQUENCE 173 AA; 20038 MW; BAF78C26B56BED31 CRC64;

Query Match 17.1%; Score 190; DB 1; Length 173;
Best Local Similarity 31.3%; Pred. No. 7e-09;
Matches 51; Conservative 32; Mismatches 54; Indels 26; Gaps 6;

QY 36 HPTLRQKAAELPLTKKEETLIAMRFLVNSQDEIAK---RYGLRSGVGLAAPQIN 91
DB 9 YPDILRLI-----LAKPYKE-----INKKIQIANDMIDTWYQEGIGLAATQVN 53

QY 92 ISKRMIAVLIPDDGSGKSYDMLVNPKIYSHVQRAYLFTGEGCLSVDDNVAGLVHRHNK 151
DB 54 IPIQLIVVNTWEO---KKONLVLIINPKIHK---KEGDISTEEGCLSIPEYQAS-IPRSNY 106

QY 152 ITIKAKDIEGNDIQRLKGYPAIVFQHEIDHLNGVMFYDHIK 194
DB 107 IQVQAVNLGDKIEIRAKSTLSICIQHEIDHLKGLKLFIDYLSK 149

Search completed: July 11, 2003, 09:42:31
Job time : 25 secs

```

GenCore version 5.1.6
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 09:31:30 ; Search time 80 Seconds
(without alignments)
548.600 Million cell updates/sec

Title: US-09-896-580A-1

Perfect score: 1112

Sequence: 1 MYELNNLFTVQLKQIKR.....KDHPIQHTDAVEVHQHHH 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_protist.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	938	84.4	183	16 Q9F4L4	Q9F4L4 staphylococ
2	536.5	48.2	183	16 Q92CX8	Q92CX8 listeria in
3	533.5	48.0	183	16 Q8Y866	Q8Y866 listeria mo
4	477	42.9	204	16 Q99X77	Q99X77 streptococ
5	462.5	41.6	203	2 Q939R9	Q939R9 streptococ
6	274.5	24.7	204	2 Q9FX63	Q9FX63 mycoplasma
7	273.5	24.6	198	16 Q9PQ25	Q9PQ25 ureaplasma
8	263	23.7	198	16 Q98PN3	Q98PN3 mycoplasma
9	226.5	20.4	224	16 Q92J77	Q92J77 rickettsia
10	218.5	19.6	188	16 Q9XN16	Q9XN16 vibrio chol
11	217.5	19.6	179	16 Q9YVH1	Q9YVH1 anabaena sp
12	215	19.3	178	2 Q93LE9	Q93LE9 leptospira
13	206.5	18.6	178	16 Q92SH6	Q92SH6 rhizobium m
14	206	18.5	273	10 Q949U8	Q949U8 arabidopsis
15	205.5	18.5	232	16 Q98RQ4	Q98RQ4 deinococcus
16	199	17.9	147	16 Q8XJ12	Q8XJ12 clostridium

17	199	17.9	284	10 Q94CZ4	Q94CZ4 oryza sativ
18	193.5	17.4	170	16 Q8UID1	Q8UID1 agrobacteri
19	191.5	17.2	177	2 Q9RQ02	Q9RQ02 zymomonas m
20	190	17.1	174	16 Q8REF0	Q8REF0 fusobacteri
21	189	17.0	150	16 Q97G95	Q97G95 clostridium
22	186.5	16.8	169	16 Q8ZLM7	Q8ZLM7 salmonella
23	184.5	16.6	159	16 Q8R9T0	Q8R9T0 thermoanaer
24	184.5	16.6	169	16 Q8ZLW9	Q8ZLW9 salmonella
25	184.5	16.6	176	16 Q98D52	Q98D52 rhizobium 1
26	184	16.5	97	2 P82590	P82590 streptococ
27	181.5	16.3	187	16 Q8YDB4	Q8YDB4 brucella me
28	180.5	16.2	155	16 Q8XJX0	Q8XJX0 clostridium
29	176	15.8	170	16 Q8ZJ79	Q8ZJ79 versinia pe
30	175	15.7	216	16 Q9K4A0	Q9K4A0 streptomyce
31	174.5	15.7	187	16 Q8YSK6	Q8YSK6 anabaena sp
32	170.5	15.3	169	16 Q8Y3B0	Q8Y3B0 raietonia s
33	169.5	15.2	173	16 Q9ABF5	Q9ABF5 caulobacter
34	169.5	15.2	218	16 Q9RD27	Q9RD27 streptomyce
35	164	14.7	403	5 Q9VG2	Q9VG2 drosophila
36	159	14.3	246	10 Q88UJ3	Q88UJ3 oryza sativ
37	154.5	13.9	162	16 Q99UQ3	Q99UQ3 staphylococ
38	152	13.7	203	16 Q8YBS4	Q8YBS4 brucella me
39	151	13.6	183	16 Q92HU7	Q92HU7 rickettsia
40	148.5	13.4	174	2 Q9ZAW6	Q9ZAW6 helicobacte
41	148.5	13.4	177	16 Q8XZJ6	Q8XZJ6 raietonia s
42	141	12.7	136	16 Q97SQ4	Q97SQ4 streptococ
43	140	12.6	136	16 Q9A0A3	Q9A0A3 streptococ
44	140	12.6	243	4 Q8WUN6	Q8WUN6 homo sapien
45	139.5	12.5	217	16 Q9FCA2	Q9FCA2 streptomyce

ALIGNMENTS

RESULT 1
Q9F4L4
ID Q9F4L4 PRELIMINARY; PRT; 183 AA.
AC Q9F4L4;
DC 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Peptide deformylase Pdf1 (pdf1 protein) (Formylmethionine deformylase homolog)
DE homolog)
GN DEFI OR PDF1 OR SAV1091 OR SA0942.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]_TaxID=158878, 158879, 1280;
RC SEQUENCE FROM N.A.
RC SPECIES=S. aureus; STRAIN=MCU29;
RA Lonetto M.A., Sylvester D.R., Warren R.L.;
RT "Staphylococcus aureus deformylase 1 encoding DNA.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=21311952; PubMed=11416146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,
Mizutani-Ti Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus.";
RL Lancet 357:1225-1240 (2001).
DR EMBL; AY007227; ABG02249.1; -
DR EMBL; AP003361; BAB57253.1; -
DR EMBL; AP003332; BAB42188.1; -

```

DR HSP; p27251; 1DEF;
DR InterPro; IPR000181; Pep_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMYLASE.
DR ProDom; PD003844; Pep_deformylase; 1.
DR TIGRFAMs; TIGR00079; Pept_deformyl; 1.
KW Complete proteome.
SQ SEQUENCE 183 AA; 20558 MW; 32A6406GA6FEABOE CRC64;

Query Match      84.4%; Score 938; DB 16; Length 183;
Best Local Similarity 98.9%; Pred. No. 1.5e-73;
Matches 181; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 25 MLTWKDIIRDGHPTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 84
DB 1 MLTWKDIIRDGHPTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60

QY 85 LAAPQINISKRMIAVLIPDDSGSKSYDMLVNPKEIVSHSVQRAYLPTGEGCLSVDDNVAG 144
DB 1 LAAPQINISKRMIAVLIPDDSGSKSYDMLVNPKEIVSHSVQRAYLPTGEGCLSVDDNVAG 120

QY 145 LVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYDHDIDKHP--LQPH 202
DB 1 LVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYDHDIDKHP--LQPH 179

QY 205 VEV 207
DB 181 VEV 183

RESULT 2
Q92CX8 PRELIMINARY; PRT; 183 AA.
ID Q92CX8 STRAIN=EGD-E / SEROVAR 1/2A;
AC Q92CX8 MEDLINE=21537279; PubMed=11679669;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-JUN-2002 (TrEMBLrel. 19, Last sequence update)
DE Hypothetical protein lin1043.
GN LIN1043.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouni F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kneft J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;
"Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596167; CAC96274.1; --
DR ListList; LIN1043; --
DR InterPro; IPR000181; Pep_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMYLASE.
DR ProDom; PD003844; Pep_deformylase; 1.
DR TIGRFAMs; TIGR00079; Pept_deformyl; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 183 AA; 20610 MW; 1D7B2637B2B73D59 CRC64;

Query Match      48.2%; Score 536.5; DB 16; Length 183;
Best Local Similarity 58.6%; Pred. No. 9.1e-39;
Matches 106; Conservative 28; Mismatches 44; Indels 3; Gaps 2;

```

```

QY 25 MLTWKDIIRDGHPTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 84
DB 1 MLTWKDIIRDGHPTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60

QY 85 LAAPQINISKRMIAVLIPDDSGSKSYDMLVNPKEIVSHSVQRAYLPTGEGCLSVDDNVAG 144
DB 1 LAAPQINISKRMIAVLIPDDSGSKSYDMLVNPKEIVSHSVQRAYLPTGEGCLSVDDNVAG 119

QY 145 LVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYDHDIDKHP--LQPH 202
DB 1 LVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYDHDIDKHP--LQPH 179

QY 203 D 203
DB 180 D 180

RESULT 3
Q8Y866 PRELIMINARY; PRT; 183 AA.
ID Q8Y866
AC Q8Y866
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein lmo1051.
GN LMO1051.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-E / SEROVAR 1/2A;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouni F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kneft J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;
"Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591977; CAC99129.1; --
DR ListList; LMO1051; --
DR InterPro; IPR000181; Pep_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMYLASE.
DR ProDom; PD003844; Pep_deformylase; 1.
DR TIGRFAMs; TIGR00079; Pept_deformyl; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 183 AA; 20643 MW; 65B2430603CDA4EF CRC64;

Query Match      48.0%; Score 533.5; DB 16; Length 183;
Best Local Similarity 57.5%; Pred. No. 1.7e-38;
Matches 104; Conservative 30; Mismatches 44; Indels 3; Gaps 2;

QY 25 MLTWKDIIRDGHPTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 84
DB 1 MLTWKDIIRDGHPTLRQKAAELEPLTKKEKETLIAMREFLVNSQDEEIAKRYGLRSGVG 60

QY 85 LAAPQINISKRMIAVLIPDDSGSKSYDMLVNPKEIVSHSVQRAYLPTGEGCLSVDDNVAG 144
DB 1 LAAPQINISKRMIAVLIPDDSGSKSYDMLVNPKEIVSHSVQRAYLPTGEGCLSVDDNVAG 119

QY 145 LVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYDHDIDKHP--LQPH 202
DB 1 LVHRHNKTIKAKDIEGNDIQLRLKGYPAIVFOHEIDHLNGVMFYDHDIDKHP--LQPH 179

QY 203 D 203
DB 180 D 180

```

QX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ReX;
RX MEDLINE=21393646; PubMed=11502510;
RA Margolis P., Hackbarth C., Lopez S., Maniar M., Wang W., Yuan Z.,
RA White R., Trias J.;
RT "Resistance of Streptococcus pneumoniae to deformylase inhibitors is
due to mutations in defB";
RL Antimicrob. Agents Chemother. 45:2432-2435(2001).
CC -!- FUNCTION: REMOVES THE FORMYL GROUP FROM THE N-TERMINAL MET OF
NEWLY SYNTHESIZED PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N-FORMYL-L-METHIONINE + H(2)O = FORMATE + L-
METHIONINE.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
DR EMBL; AV014509; AAK13238.1; -;
DR InterPro; IPR000181; Pep deformylase.
DR Pfam; PF01327; Pep deformylase; 1.
DR PRINTS; PR01576; PDSEFORMYLASE.
DR ProDom; PD003844; Pep_deformylase; 1.
DR TIGRfams; TIGR00079; pep deformyl; 1.
KW Hydrolyase; Protein biosynthesis; Zinc.
SQ SEQUENCE 203 AA; 22692 MW; E332956982A67161 CRC64;

Query Match 41.6%; Score 462.5; DB 2; Length 203;
Best Local Similarity 50.8%; Pred. No. 2.7e-32;
Matches 95; Conservative 37; Mismatches 48; Indels 7; Gaps 2;

QY 18 KIRKVOYMLTKDIIRDGHPTLRQKAELPLTKKEKETLIAMREFLVNSODEIAKY 77
Db :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
6 RTKAHLDMNDIIREGNETLTVAEEVTFPLSDQEIILLGECMQFLKHSQDPVMAEK 65

QY 78 GLRGSGVLGAAPQINISKEIMAVLIPD-----DGSGKSVDY--MLNPKIVSHSVQEA 130
Db |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
66 GLRGSGVLGAAPQLDISRIIAVLNVPIVEGETPQEADIEALWNPVKIVSHSVQDAALG 125

QY 131 TGEGCLSYDDNVAGLVHRHNKITIKAKDIGNDIQLRKKGYPAPVFOHEIDLNGVMFYD 190
Db |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
126 EGEGCLSYDVNRVPGVVVRHARVTVDYFDKGEKHRIKLKGVNSIVVQHEIDHNGVMFYD 185

QY 191 HDKDHP 197
Db :|::|
186 RINEXDP 192

RESULT 6
QSKX63 PRELIMINARY; PRT; 204 AA.
AC Q9FK63
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polypeptide deformylase (EC 3.5.1.31) (PDF) (Formylmethionine
deformylase).
GN DEF OR DEF.L.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A5969;
RA Skamrov A., Feoktistova E., Goldman M., Bebealashvili R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DDJB databases.
CC -!- FUNCTION: REMOVES THE FORMYL GROUP FROM THE N-TERMINAL MET OF
NEWLY SYNTHESIZED PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N-FORMYL-L-METHIONINE + H(2)O = FORMATE + L-
METHIONINE.
CC -!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
DR EMBL; L35043; AAF36760.1; -;
DR HSSP; P27251; 2DEF.

QY 203 D 203
Db 180 D 180

RESULT 4
Q9XY7 PRELIMINARY; PRT; 204 AA.
AC Q99X17
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative polypeptide deformylase.
GN DEF OR SPY1956.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=1296299;
RA Ferretti U.J., McShan W.N., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvarov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren O., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006618; AAK34651.1; -;
DR HSSP; P27251; 2DEF.
DR InterPro; IPR000181; Pep deformylase.
DR Pfam; PF01327; Pep deformylase; 1.
DR PRINTS; PR01576; PDEFORMYLASE.
DR ProDom; PD003844; Pep_deformylase; 1.
DR TIGRfams; TIGR00079; pep_deformyl; 1.
KW Complete proteome.
SQ SEQUENCE 204 AA; 22862 MW; 89F8E9D94DC05 CRC64;

Query Match 42.9%; Score 477; DB 16; Length 204;
Best Local Similarity 52.6%; Pred. No. 1.5e-33;
Matches 101; Conservative 33; Mismatches 50; Indels 8; Gaps 2;

QY 16 QIKRKVOYMLTKDIIRDGHPTLRQKAELPLTKKEKETLIAMREFLVNSODEIAK 75
Db :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
4 QDKLIKPSHLITMDIIREGNETLTVAEEVTFPLSDQEIILLGECMQFLKHSQDPVMAE 63

QY 76 RYGLRSVGVLGAAPQINISKEIMAVLIPD---DGSGK-----SYDVLNPKIVSHSVQEA 127
Db :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
64 KLGRAGVGLGAAPQIDVSRKIIVNLFPDKGNPPKAYSWQEVLLNPKIVSHSVQQA 123

QY 128 YLFTGEGCLSYDDNVAGLVHRHNKITIKAKDIGNDIQLRKKGYPAPVFOHEIDLNGVM 187
Db :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
124 ALSDGEGCLSYDVNRVPGVVVRHARVTVDYFDKGEQQHQRIKLKGVNAIVVQHEIDHNGVL 183

QY 188 FYDHIDKDELQ 199
Db :|::|:
184 FYDIRNAKNFE 195

RESULT 5
Q939R9 PRELIMINARY; PRT; 203 AA.
AC Q939R9
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polypeptide deformylase (EC 3.5.1.31) (PDF) (Formylmethionine
deformylase).
GN DEF OR DEF.B.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.

RESULT	Q92J17	PRELIMINARY;	ID	Q92J17	PRT;	224 AA.
AC	Q92J17		AC	Q92J17		
DT	01-DEC-2001	(TRENBSrel. 19, Created)	DT	01-DEC-2001	(TRENBSrel. 19, Last sequence update)	
DT	01-DEC-2001	(TRENBSrel. 19, Last sequence update)	DT	01-DEC-2001	(TRENBSrel. 19, Last sequence update)	
DT	01-JUN-2002	(TRENBSrel. 21, Last annotation update)	DT	01-JUN-2002	(TRENBSrel. 21, Last annotation update)	
DE		Polypeptide deformylase.	DE		Polypeptide deformylase.	
GN	DEF3 OR KC0080.		GN	DEF3 OR KC0080.		
OS	Rickettsia conorii.		OS	Rickettsia conorii.		
OC	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;		OC	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;		
OC	Rickettsiaceae; Rickettsia.		OC	Rickettsiaceae; Rickettsia.		
OC	NCBI Taxid=781;		OC	NCBI Taxid=781;		

```

[1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Coseart P., Weissbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098 (2001).
DR EMBL; AF008575; AA02618.1; -.
DR InterPro; IPR000181; Pep_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMLASE.
DR ProDom; PD003844; Pep_deformylase; 1.
KW Complete proteome.
SQ SEQUENCE 224 AA; 25539 MW; 74ED5CB61C16237F CRC64;

Query Match      20.4%; Score 226.5; DB 16; Length 224;
Best Local Similarity 34.1%; Pred. No. 8.6e-12;
Matches 72; Conservative 29; Mismatches 73; Indels 37; Gaps 9;

QY 12 IQLKQIKIRKQVQMLTMDIIRDGHTLRQKAAELELPKTEE-KETLIAMRFLVNSQD 70
DB 12 IQLKQIKIRKQVQMLTMDIIRDGHTLRQKAAELELPKTEE-KETLIAMRFLVNSQD 70
QY 71 BEAKRYGLRSGVGLAALPQINISKRMIAVLIPDDGSKSY--DYMLVNPKIVSHSVQEAY 128
DB 71 BEAKRYGLRSGVGLAALPQINISKRMIAVLIPDDGSKSY--DYMLVNPKIVSHSVQEAY 128
QY 80 ENCA-----GLAARQIGISKIILFAVHEDAEKXKHPDLKDTMPTKI--WNPST 128
DB 80 ENCA-----GLAARQIGISKIILFAVHEDAEKXKHPDLKDTMPTKI--WNPST 128
QY 129 LPTG-----RCGLSVDDNVAGLVHRHNKTIKAKDIEGNDIQRLKGYPAIVFQHEID 181
DB 129 LPTG-----RCGLSVDDNVAGLVHRHNKTIKAKDIEGNDIQRLKGYPAIVFQHEID 181
QY 129 KPIGIDKHEDYEGCFSV-ENATGPFARFKKHVHAYDINGNQIQIAGFLARVQIHEID 187
DB 129 KPIGIDKHEDYEGCFSV-ENATGPFARFKKHVHAYDINGNQIQIAGFLARVQIHEID 187
QY 182 HLGKGVWFYDHI-----DKQHPLOPHDTAVS 206
DB 182 HLGKGVWFYDHI-----DKQHPLOPHDTAVS 206
QY 188 HLGKGVFLDYVAPKIMTKKEEYLEMRKKAME 218
DB 188 HLGKGVFLDYVAPKIMTKKEEYLEMRKKAME 218

RESULT 10
QYKNI6 PRELIMINARY; PRT; 168 AA.
AC Q9KNI6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Polypeptide deformylase.
GN VCA0150.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Bisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Eimolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.D., Venter J.C.,
RA Fraser C.N.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483 (2000).
DR EMBL; AE004356; AAF96063.1; -.
DR HSSP; P27251; IDFF.
DR TIGR; VCA0150; -.
DR InterPro; IPR000181; Pep_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMLASE.
DR ProDom; PD003844; Pep_deformylase; 1.
DR TIGRFAMs; TIGR00079; pep_deformyl; 1.
KW Complete proteome.

```

```

SQ SEQUENCE 168 AA; 18671 MW; F51EDDB3B6C14C59 CRC64;

Query Match      19.6%; Score 218.5; DB 16; Length 168;
Best Local Similarity 31.4%; Pred. No. 2.9e-11;
Matches 58; Conservative 39; Mismatches 65; Indels 23; Gaps 6;

QY 26 LTWKDIIRDGHTLRQKAAELELPKTEEKTLIAMRFLVNSQDEEIAKRYGLRSGVGL 85
DB 1 MAYLEILTAPDPRFVOSKQV-----TDVASVOTLI---DLDLTLATDNGIGL 47
QY 86 AAPQINISKRMIAVLIPDDGSKSYDYMLVNPKIVSHSVQEAYLPTGEGCLSVDDNVAGL 145
DB 86 AAPQINISKRMIAVLIPDDGSKSYDYMLVNPKIVSHSVQEAYLPTGEGCLSVDDNVAGL 145
QY 48 AARQVG---REEAIVIDLSDNRDQPLVILNPKVSGSKEM---GOEGCLSVFDIYAD- 100
DB 48 AARQVG---REEAIVIDLSDNRDQPLVILNPKVSGSKEM---GOEGCLSVFDIYAD- 100
QY 146 VHRHNKTIKAKDIEGNDIQRLKGYPAIVFQHEIDHLNGVNFYDHDKHPLOPHDTAV 205
DB 101 VERYTSVVVEALDREGKPLRTSDFLAIVMQHEIDLHSGNLFIDYLS---PLKQQWAMK 157
QY 206 EVHOH 210
DB 158 KVKKH 162

RESULT 11
QYVHL PRELIMINARY; PRT; 179 AA.
AC Q8YVHL;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Polypeptide deformylase.
GN ALI2007.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213 (2001).
DR EMBL; AP003587; BAB73706.1; -.
DR InterPro; IPR000181; Pep_deformylase.
DR Pfam; PF01327; Pep_deformylase; 1.
DR PRINTS; PR01576; PDEFORMLASE.
DR ProDom; PD003844; Pep_deformylase; 1.
DR TIGRFAMs; TIGR00079; pep_deformyl; 1.
KW Complete proteome.
SQ SEQUENCE 179 AA; 20136 MW; E17DBCB9B0E473D2 CRC64;

Query Match      19.6%; Score 217.5; DB 16; Length 179;
Best Local Similarity 33.1%; Pred. No. 3.9e-11;
Matches 57; Conservative 30; Mismatches 60; Indels 25; Gaps 5;

QY 31 IIRDGHTLRQKAAELELPKTEEKTLIAMRFLVNSQDEEIAKRYGLRSGVGLAAPQI 90
DB 8 IIOGNPTLRQKAWVE-----NIHDATIQLLIDDLIATVAKANGVGIASPOV 55
QY 91 NISKRMIAVLIPDDGSKSYDY-----MLVNPKIVSHSVQEAYLPTGEGCLSVDDNVA 143
DB 56 AQSRLFLIVA---SRPNRYPHAPENPTAMINPKIVGHSTE--IVEGMEGCLSV-PGIR 109
QY 144 GLVHRHNKTIKAKDIEGNDIQRLKGYPAIVFQHEIDHLNGVNFYDHDKID 195
DB 110 GLVFRQALEVEYTDYRGNLQKQTLTDFVARIFQHEFDHLDGVLFDIVRESN 161

RESULT 12

```

RT	"Analysis of the chromosome sequence of the legume symbiont	
RT	Sinorhizobium meliloti strain 1021.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).	
DR	EMBL; AL591783; CAC41856.1;	
DR	InterPro; IPR000181; Pep_deformylase.	
DR	Pfam; PF01327; Pep_deformylase; 1.	
DR	PRINTS; PR01576; PDEFORMYLASE.	
DR	ProDom; PD003844; Pep_deformylase; 1.	
DR	TIGRFAMs; TIGR00079; pept_deformyl, 1.	
KW	Hydrolase; Complete proteome.	
SEQUENCE	178 AA; 19893 MW; 98F83DC594E38799 CRC64;	
Query Match 18.6%; Score 206.5; DB 16; Length 178;		
Best Local Similarity 34.9%; Pred. No. 3.5e-10;		
Matches 59; Conservative 28; Mismatches 65; Indels 17; Gaps 6;		
QY	26	LTWKQIIRDGHPTLRQAAELELPITKEKETLTAMREFLVNSQDEETAKRYGLRSVGVL 85
DB	5	MTTKPILPDPVLVRQVSTPV-----ETIDATIRRLA---DMLLETMYD-APGIGL 51
QY	86	AAPQINISKRIMAVLI PDDGSGKSYDMLVNPKNVSHSVQEAFLPTGSGKSDVDNVAGL 145
DB	52	AAQLQGVKRLVLDVTKGSEKQ-PLVFIKPKVRSERSVYE--EGCLSIDPYTAE- 107
QY	146	VHRHNKITIKAKDIGNDICLRKGYPAIVFQHEIDHNGVMFYDHDK 194
DB	108	VERPAATVEYVDREGKEQAVEAGLLATCQLEIDHNGVLFIDVISK 156
RESULT 14		
Q949U8	PRELIMINARY; PRT; 273 AA.	
ID	Q949U8	
AC	Q949U8;	
DT	01-DEC-2001 (TrEMBLrel. 19, Created)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)	
DE	Hypothetical 30.6 kDa protein.	
OS	Arabidopsis thaliana (Mouse-ear cress).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC	eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
OX	NCBI_TaxID=3702;	
RP	[1]	
RP	SEQUENCE FROM N.A.	
RA	Yamada K., Liu S.X., Pham P.K., Banno F., Dale J.M.,	
RA	Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,	
RA	Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,	
RA	Carninci P., Chen H., Cheuk R.K., Hayaishizaki Y., Ishida J., Jones T.,	
RA	Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Kossena E., Lam B.,	
RA	Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,	
RA	Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,	
RA	Shinozaki K., Davis R.W., Ecker J.R., Theologis A.	
RT	Full Length cDNA of gene T1EN1_150/AT5G14560 (GI:7573315)";	
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AY050879; AK92816.1; -	
DR	InterPro; IPR00181; Pep_deformylase.	
DR	Pfam; PF01327; Pep_deformylase; 1.	
DR	PRINTS; PR01576; PDEFORMYLASE	
DR	ProDom; PD003844; Pep_deformylase; 1.	
DR	TIGRFAMs; TIGR00079; pept_deformyl, 1.	
KW	Hypothetical protein.	
SEQUENCE	273 AA; 30610 MW; 25CDA90ED6D9603E CRC64;	
Query Match 18.5%; Score 206; DB 10; Length 273;		
Best Local Similarity 32.2%; Pred. No. 6.9e-10;		
Matches 55; Conservative 33; Mismatches 65; Indels 18; Gaps 6;		
QY	22	VOYMTMKDITRDGHTLRQAAELELPITKEKETLTAMREFLVNSQDEETAKRYGLRS 81
DB	75	VQEPFLK-IVEYDPDLRAKNKRDI-FDNLKLVDMFVNKYT-----D 120
QY	82	GVGLAAPQINISKRIMAVLPDGGSGKSYDMLVNPKNVSHSVQEAFLPTGEGCLSDVDN 141

Search completed: July 11, 2003, 09:40:42
Job time : 83 secs